DOCUMENT RESUME

ED 148 807 95 SP 012 079

AUTHOR	Fitz-Gibbon, Carol Taylor
TITLE	An Analysis of the Literature on Cross-Age Tutoring. CSE Report on Tutoring, No. 5.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Nov 77
NOTE	119p.; For related documents, see SP012 074-080
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	*Cross Age Teaching; *Educational Research; *Elementary Secondary Education; *Literature Reviews; *Tutorial Programs

ABSTRACT

This document examines eighteen studies on the subject of cross-grade tutoring, the studies being classified according to the manner in which the tutoring task is selected--whether based on the needs of the teacher, the tutor, or the tutee. Presentation of the studies follows the following sequence: (1) a description and analysis of experiments run for short periods of time by persons external to the schools; (2) projects run by school personnel ("field studies"); and (3) "programmed tutoring," in which the tutor administers a predesigned program of instruction, and which is usually conducted for the benefit of the tutee. Analyses of the reports concentrates mainly on cognitive learning, though a brief survey of noncognitive outcomes is included. (MJB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

Center for the Study of Evaluation

UCLA Graduate School of Education
Los Angeles, California 90024

CSE
CSECSE
CSECSECSE
CSECSE
CSE

ED148807

SP012 079

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Center for the Study of Evaluation

The Center for the Study of Evaluation (CSE) is a university-based research and development center established in 1966 by the Department of Health, Education, and Welfare. The mission of the Center has been to conduct research, development, training, and dissemination activities for the purpose of improving the planning and evaluation of education and other social action programs. More specifically, CSE activities include: guidance for educators and researchers in the areas of tests and measures and in educational decision making; development of easy to use evaluation procedures and materials in the form of manuals, workshops and kits; and evaluation of projects for federal, state, local, and university agencies requesting technical assistance in evaluation. The work of CSE during its first decade has clarified the ways in which systematic evaluation can be helpful in the improvement of many kinds of programs.

Information about CSE and its publications may be obtained by writing to:

Director, Public Information
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

CSE Report on Tutoring #5:
An Analysis of the Literature on Cross-Age Tutoring

Carol Taylor Fitz-Gibbon

Center for the Study of Evaluation
UCLA Graduate School of Education

November 1977

This work was performed in part pursuant to a grant [contract] with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official National Institute of Education position or policy.

TABLE OF CONTENTS

PREFACE i

ACKNOWLEDGEMENTS iv

CHAPTERS

I. Introduction 1

- A Classification Scheme for Tutoring Projects 1
- The Organization of This Report 4

II. Cognitive Outcomes in Experiments and Field Studies 8

Experiments in School Settings 8

- Being a Tutor vs. Remaining in Class 8
- Effectiveness of Low Achieving Students as Tutors 11
- Being a Tutor vs. Studying Alone 16
- Possible Processes Involved in Promoting Learning in Tutors 18
- Being a Tutee vs. Being in a Class 21
- Misleading Claims 23
- Caution re Standardized Tests 26

Field Studies with Secondary School Tutors 27

- A Youth Tutoring Youth Project 27
- Three Cross-School Tutoring Projects 34
- A Within-Secondary-School LbT Project 36
- A Project that Nearly Failed 40
- Ad Hoc Remediation vs. Full Sequence Instruction 42

Field Studies with Elementary School Tutors 43

- Two Intact Class Tutoring Projects 43
- More on Ad Hoc Remediation 47
- A Learning-by-Tutoring, Full Sequence Instruction Project 50
- Tutors Gain When the Task Matches Their Needs 53
- More on the Issue of Standardized Tests 54

Summary of Cognitive Outcomes 58

III. Cognitive Outcomes in Programmed Tutoring 61

- Effects on Tutors of Administering Programmed Tutoring 61
- Programmed vs. "Directed" Tutoring for Beginning Reading 65
- Summary of Programmed Tutoring 83

IV. Non-Cognitive Outcomes of Tutoring 84

- Summary of Non-Cognitive Outcomes 91

APPENDICES

A. Classifications of Tutoring Projects 92
B. Reading About Tutoring 98

REFERENCES

LIST OF FIGURES

Figure 1. Procedures for three kinds of tutoring projects 3

Figure 2. Characteristics of major studies examined 7

Figure 3. Ninth grade posttest results 10

Figure 4. Fourth grade posttest results 12

Figure 5. Scores of experimental and control groups of the Scrambled Fractions posttest 13

Figure 6. Allen and Feldman results for tutors 17

Figure 7. Allen and Feldman results for tutees 17

Figure 8. Hamblin and Hamblin results 24

Figure 9. Myers, Travers, and Sanford results 63

Figure 10. Average responses to the question "Would you like to have tutoring as an elective one quarter?" 87

Figure 11. Responses of tutors and non-tutors to tutoring 88

Figure 12. Comparison of three classification schemes 96

Figure 13. Examples of projects classified by the amount of structuring of the tutoring task 97

LIST OF TABLES

Table 1. Cloward's Results for Tutees 30

Table 2. Cloward's Results for Tutors 32

Table 3. Hoffmeister's Results for Tutors and Tutees 39

Table 4. First Grade Stanford Reading Test Results 45

Table 5. Second Grade Stanford Reading Test Results 46

Table 6. Niedermeyer and Ellis Results Regarding Effectiveness of Tutors 48

Table 7. Means and F-test Between the Post-test Scores of Control and Experimental Tutors 51

Table 8. Means and F-test Between the Post-test Scores of Control and Experimental Tutees 52

Table 9. Ellson, Harris and Barber Results 69

Table 10. Ellson, Harris and Barber Analyses of Covariance 70

Table 11. Ronshausen's Summary of Math Concepts Tests 77

Table 12. Ronshausen's Summary of Math Computational Skills 78

PREFACE

The National Institute of Education (NIE) is the agency of the Federal government that is specifically funded to conduct research in education. When the Education Amendments of 1974 were enacted, NIE was directed to prepare studies of compensatory education in order to assist Congress in making decisions regarding the future of efforts to improve the education of disadvantaged youth. The studies were mandated with a view to the reauthorization deliberations for Title I of the Elementary and Secondary Education Act (ESEA), scheduled to take place in 1977.

In connection with these studies, contracts were let for surveys of existing compensatory education programs, analyses of existing data on program effectiveness, research into the implications of alternative criteria for fund-allocations, and research on the administration of compensatory education. In addition, NIE requested "a systematic and interdisciplinary effort to explore alternatives to current practices in education." Four contracts were let for "Alternative Designs to Contrast with Current Practice in Compensatory Education." Each contractor was to produce "a plan of action which is deemed superior to current practice and not prohibitive in cost of implementation."

One of the contracts went to the Center for the Study of Evaluation for a proposal concerned with cross-age tutoring as a means of enhancing the learning and motivation of the tutors (in contrast to current practice in which the learning of the tutee is the primary focus).

An interdisciplinary committee of ten persons was established to meet regularly at UCLA. In addition to working with this committee, project staff undertook to develop an examination of current school practice by means of site visits and detailed questionnaires in schools conducting cross-age tutoring projects. Because the persons who ultimately implement innovations are the instructional staff, it was important to obtain their reactions to the proposed projects. Presentations were made at faculty meetings in inner city schools, followed up with questionnaires, and a small conference was held at which it was possible to obtain more detailed feedback.

Following these activities, an interim report was submitted to nine reviewers selected by NIE. Committee members and reviewers subsequently met together for two days of round-table discussions.

The production of a single final report would have been an inappropriate way of presenting the body of information that project staff had at hand following these activities. For example, much information that would be of interest to teachers and parents planning or running tutoring projects would not be of interest to policy makers and vice versa. Furthermore, although the position had been reached that one particular kind of tutoring project--the Learning-Tutoring Cycle (LTC)--should be recommended for widespread implementation in Title I projects, there was much to be said about the planning of tutoring projects in general and about highly innovative projects which, although not as immediately feasible as the Learning-Tutoring Cycle, could have greater impact if implemented.

In view of this need to speak to various audiences about diverse topics such as theories, research, policy, practical planning and innovative ideas, seven separately bound volumes were prepared: The CSE Reports on Tutoring.

Report A.	The Learning-Tutoring Cycle: Overview
Report #1	Setting up and Evaluating Tutoring Projects
Report #2	Tutoring and ESEA Title I
Report #3	A Survey of Tutoring Projects
Report #4	Tutoring: Some New Ideas
Report #5	An Analysis of the Literature on Cross-Age Tutoring
Report #6	Tutoring and Social Psychology: A Theoretical Analysis

These seven reports provide an information base and a rationale for actions at both federal and local levels.

At the local level, the six numbered reports represent a resource for project planning. Three reports--the Survey (#3), the Literature report (#5), and the Social Psychology report (#6)--bring together information from a wide range of sources to provide background knowledge concerning current practice, the perceptions and recommendations of practitioners, past projects, research, and relevant theories.

This background knowledge can inform the design of tutoring projects as well as provide a basis for judging the likelihood of success of such projects. For example, the Social Psychology report demonstrates that many of the effects of tutoring widely noted by practitioners, but not yet established by research on tutoring, have their parallel in recognized effects in the field of Social Psychology. Social-Psychological theories thus suggest variables that can be assessed in research on tutoring and lend support to the perceptions of practitioners.

The three other numbered reports apply the background information to practical school planning concerns. Report #1, through a presentation of step-by-step planning questions, deals with the process of setting up and evaluating tutoring projects in general. Report #2 introduces the ESEA Title I legislation and considers the planning steps, as per Report #1, in relation to the design of an LTC project in conformity with Title I regulations. Report #4 deals with extensions of the idea of learning by tutoring to more ambitious projects, representing more substantial changes in schooling than are involved in the LTC project. Report #4 also documents the reactions of instructional staff and some Title I parents to the Learning-Tutoring Cycle and other proposed kinds of projects.

For the national level, Report A provides an overview of the project and recommendations for action. The actions recommended are designed to put the ideas of this study to the test, to see if the Learning-Tutoring Cycle can indeed significantly improve the educational attainment of disadvantaged students.

ACKNOWLEDGEMENTS

This study has benefitted immeasurably from the thoughtful contributions of many individuals. Members of the committee represented diverse backgrounds and professional experiences, and their deliberations were invariably stimulating and challenging.

Members of the Committee

- Harry F. Silberman (Co-Principal Investigator)
  Professor of Education
  UCLA

- Marvin C. Alkin
  Professor of Education
  UCLA

- Ron Apperson
  Legal Adviser
  Los Angeles Unified School District

- Dorothy Bradshaw
  General Secondary Curriculum Consultant
  Compton Unified School District

- Erick Lindman
  Professor of Education
  UCLA

- Neil Malamuth
  Professor of Psychology
  UCLA

- Gerald Newmark
  Co-Director
  Educational Communications Corp.

- Floraline Stevens
  Coordinator, Research and Evaluation
  Los Angeles Unified School District

- Richards Williams
  Research Associate
  Center for the Study of Evaluation

We were fortunate in having two ex-officio members of the committee who attended meetings because of their long-standing interest in Tutoring: Elbert Ebersole, formerly Principal, Soto Street Elementary School and now a developer of materials for cross-age tutoring in reading; and Arthur Elliott, a professor on sabbatical leave from Simon Fraser University, Canada. Their contributions were of great value to the project staff and to the committee.

The round table discussions with members of the NIE review team and the receipt of their carefully considered written comments on the interim report led to significant shifts of emphasis in the project, and the author is most grateful for their substantial and thoughtful comments.

Members of the NIE Review Panel

- Bill Anderson
  Coordinator
  National Coalition of Title I Parents

- Larry Barber
  Asst. Supt., Research, Development & Evaluation
  Eugene Public Schools

- Sophie Bloom
  Reading Coordinator
  Chicago Public Schools

- Linda Glendening
  Research Associate, Measurement & Methodology Division
  National Institute of Education

- Philip Hawkins
  Special Assistant for Planning
  Michigan Department of Education

- Gerry Hendrickson
  Education Program Specialist
  Department of Health, Education & Welfare

- Gwyn Jones-Davis
  Director
  Antioch Communion University

- Robert Lipscomb
  President
  Alabama Education Association

- David Weikert
  President
  High-Scope Education Research Foundation

Special thanks are due to the many persons working in schools who met with project staff, conducted site visits or responded to the lengthy questionnaire. Likewise, the teachers and administrators who attended the conference or gave us valuable feedback at their staff meetings are offered sincere appreciation for their contributions.

Richard Mallory of the National Education Association was so kind as to comment extensively and most helpfully on a draft report. The incredible efficiency of Donna Anderson Cuvelier and her great interest in the project provided indispensable daily support. For contributions and guidance throughout the study, the author thanks most particularly Harry F. Silberman and Jack Schwille.

CHAPTER I. INTRODUCTION

Since there are a number of literature reviews available for the topic of cross-age tutoring (Devin-Sheehan, Feldman and Allen, 1976; Bloom, 1976; Klaus, 1975; Rosenshine and Furst, 1969), it is important to specify the distinctive features of the present analysis. In this analysis a few studies have been selected for examination in detail. The level of detail is such that when the conclusions of the original authors are rejected or questioned (e.g., as in Hamblin and Hamblin, 1972; Rosenshine and Furst, 1969, and Niedermeyer and Ellis, 1971), readers will find sufficient information in this report to form their own judgments concerning the critique offered.

The fairly detailed examination of selected studies also serves to illustrate various points concerning both tutoring as an instructional procedure and concerning the methodology of research and evaluation. The description of a variety of projects and experiments should prove helpful to school district personnel and instructional staff planning studies of their own tutoring projects.

A Classification Scheme for Tutoring Projects

What might one expect to find in a school reported to have a "tutoring project"? A little experience reveals that a wide variety of organizational changes may have occurred under the general rubric of a "tutoring project." In some projects two or three older students may work in an elementary classroom, assisting the classroom teacher in various ways. In other projects a class of secondary students may be brought to an elementary school and supervised by their own teacher while tutoring children pulled out of their regular classes to receive one-to-one instruction. Different again is a project in which older students act as counselors to younger students, or a project in which an entire upper grade class meets regularly with an entire lower grade classroom for one-to-one work with a set of materials designed to teach reading skills. To consider all these projects as likely to produce the same kind of results is to ignore their substantial differences. Some classification and analysis of characteristics must be attempted.

There are many characteristics that differ even among the few examples just cited--grade levels, pull-out vs. intact class, cognitive vs. affective goals, amount of supervision, location, degree of integration of tutoring with the regular school program, etc. On what characteristics should a classification be based, a classification that will differentiate between projects likely to have different effects? It is suggested here that the crucial question to ask about a tutoring project concerns the nature of the tutoring task: What is to occur between tutor and tutee? How is the task selected and defined each day and throughout the project? It is on the basis of how the tutoring task is selected that projects will be classified.

If the tutoring task is selected on the basis of the teacher's need for various kinds of assistance in running a classroom, then the project is a Teacher Aide project. If the tutoring task is selected on the basis of tutees' needs, such as a need for instruction in a particular area, or for remedial help, or for counseling or friendship, then the project will be called a Tutorial Service project. If the tutoring task is selected to meet identified needs of tutors, such as their need to learn mathematics, or responsibility or empathy, then the project will be termed a Learning-by-Tutoring project. Figure 1 illustrates how the procedures differ among the three kinds of projects.

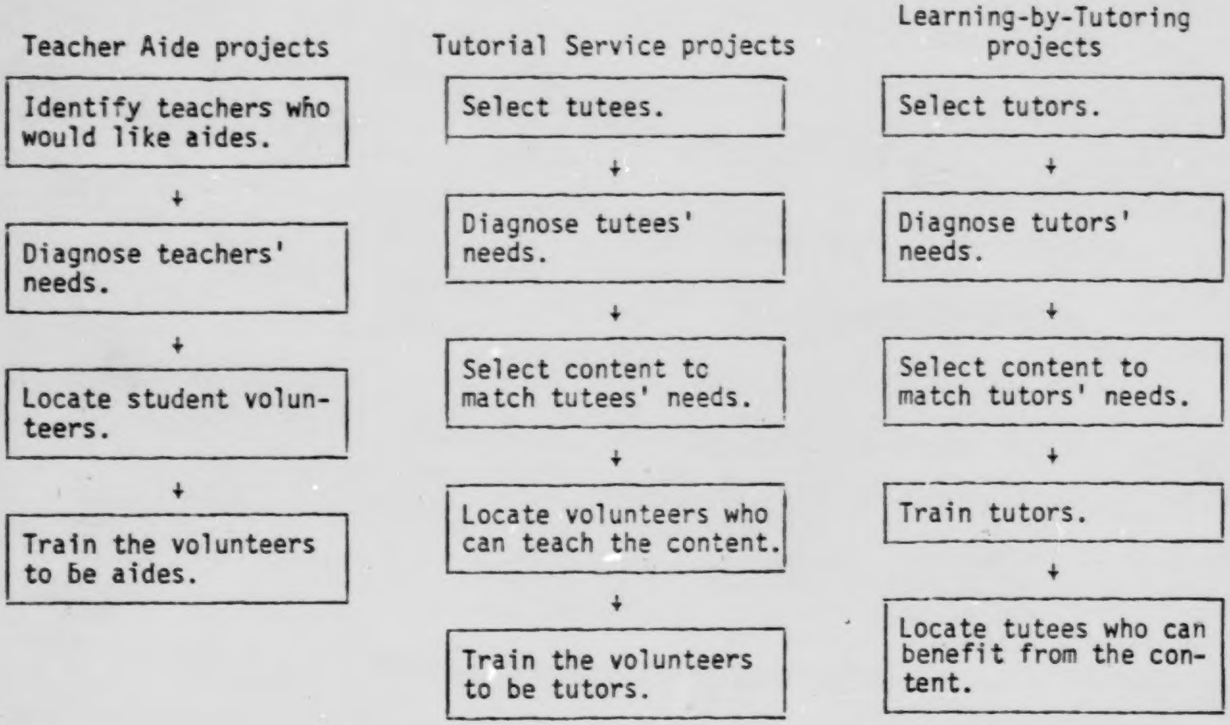
Teacher Aide projects
Identify teachers who would like aides.
Diagnose teachers' needs.
Locate student volunteers.
Train the volunteers to be aides.
Tutorial Service projects
Select tutees.
Diagnose tutees' needs.
Select content to match tutees' needs.
Locate volunteers who can teach the content.
Train the volunteers to be tutors.
Learning-by-Tutoring projects
Select tutors.
Diagnose tutors' needs.
Select content to match tutors' needs.
Train tutors.
Locate tutees who can benefit from the content.

Figure 1. Procedures for three kinds of tutoring projects

The distinction between Tutorial Service (TS) and Learning by Tutoring (LbT) projects needs emphasis. In a TS project, tutees needs are diagnosed, and then tutors are sought who can teach the content that tutees need to learn (or meet tutees' other needs if the project is not one concerned with learning in the academic sense). In LbT projects the needs of tutors are diagnosed, and the tutoring task will be defined by those needs. Then tutees are located for whom the instruction (or other activities of the task) will be suitable.

Although it is not an established finding, there is a good deal of anecdotal evidence to suggest that tutors will lose interest if their tutees are not clearly benefitting from the tutors' activities. Thus in LbT projects, it is probably vital to ensure that, although the project was designed to enable tutors to learn by tutoring, tutees receive appropriate instructions and benefit from it. Thus LbT projects require that both tutors and tutees benefit. This is ensured by the selection of appropriate tutees rather than be adjustment in the tutoring task. In Appendix A, the classification of tutoring projects shown in Figure 1 is contrasted with two other classifications that have been proposed.

The Organization of This Report

In Figure 2, the studies that are discussed in most detail in this report are listed along with a number of their basic characteristics: kind of project, size and duration, relationship to the regular school program, and population characteristics. Since the studies are listed in the order in which they are considered in the report, it can be seen from Figure 2 that the report begins with a description and analysis of several experiments run for short periods of time by persons external to the schools. These studies have the virtue of data collected in the framework of a good experimental design involving control groups formed by random assignment.

The next set of studies examined ("field studies") includes only projects run by school personnel. These studies generally lack readily interpretable data since the information, if any, tends to be standardized test scores for participants only. These scores may be compared with general expectations but do not provide strong evidence for the effects of tutoring. (Readers not familiar with Campbell and Stanley, 1966, are advised to read that influential analysis of experimental designs.) A major virtue of the field studies is that their existence demonstrates the possibility of large scale projects.

A third category of study examined is the "programmed tutoring" project. This kind of tutoring places the tutor in the role of administering a predesigned "program" of instruction and is usually conducted for the benefit of the tutee, i.e., a Tutorial Service project.

The analyses of experiments, field studies, and programmed tutoring projects constitute Chapters II and III. In the course of these chapters, a number of issues are discussed. These issues are noted in the table of contents which therefore serves to some extent the function of an index.

In Chapters II and III, the analyses concentrates on cognitive learning outcomes. A brief survey of "non-cognitive" outcomes of tutoring projects is provided in Chapter IV. This imbalance--many more pages devoted to cognitive outcomes than to non-cognitive outcomes--is not present because cognitive outcomes are necessarily seen as more important than other outcomes. One cause of the imbalance is probably that cognitive outcomes are more often measured and are perhaps more accurately measurable.

The relative importance of the cognitive versus "affective" impact of tutoring is not considered a fruitful area for discussion, since such discussion would be unlikely to affect practice. Local personnel planning projects will have their own orientations and will presumably implement their projects in conformity with these and with the wishes of the community. However, since so many of the following pages will be concerned with cognitive outcomes, an elegant quotation from Jerome Bruner is presented here to conclude this introduction and to redress the balance a little:

> ...when we read Urie Bronfenbrenner's perceptive analysis of the isolating tendency within American society to grade and segregate by age, one realizes that the idea of each teaching another may indeed be a revolutionary step toward maintaining community in a society where the forces of urban organization, of mass production, and indeed of mass education are all centrifugal (p. 62).
>
> The contemporary effort to reconstruct our schools is, I think, part of the general effort to reduce isolation, to re-establish mutuality and exchange. Sharing one's skills and knowledge with others, being teacher as well as learner are efforts to that end.
>
> Perhaps superintendents, principals, and school boards may have to be lured with pragmatic rationale. It is not the right reason for adopting this plan. It is a bonus. The real reason should be to get us an inch on the way toward making the helper and the helped the universal unit of exchange within a culture that continues to produce lonely crowds, lonelier than ever (p. 63).

TUTORING STUDY	KIND OF PROJECT			GRADE LEVELS, SIZE AND DURATION					RELATIONSHIP TO REGULAR SCHOOL			
				TUTORS		TUTEES						
(page no.)	TA	TS	LbT	#	Grade	#	Grade	D	Run by school personnel?	In regular school hours?	CONTROL GROUP?	"DISADVANTAGED" POPULATION?
Fitz-Gibbon (8)			✓	40	9th	68	4th	3 wks	no	yes	yes	yes
Allen and Feldman (16)			✓	10	5th	10	3rd	2 wks	no	no	yes	no
Bausell, Moody & Walzl (21)		✓			adult	60	4th/5th	1 da	no	yes	yes	no
Hamblin & Hamblin (23)		✓		?	preschool & teenage	32	preschool	8 wks	no	yes	(yes)	yes
Snapp, Oakland & Williams (26)			(✓)	60	5th/6th	60	2nd/3rd	8 wks	no	no	yes	yes
Cloward (27)		✓		240	10th/11th	410	4th/5th	>1 yr	partially	no	yes	yes
Ontario/Montclair California (34)			✓		secondary		elementary	>1 yr	yes	yes	(no)	some projects
Madison, Wisconsin (35)	✓				secondary		elementary	>1 yr	yes	yes	no	?
Paducah, Kentucky (35)	✓			320	secondary		elementary	>1 yr	yes	yes	no	yes
Hoffmeister (36)			✓	41	9th	40	7th	>1 yr	yes	yes	no	yes
Dreyer (40)		✓			7th		1st	1 yr?	yes, with help	yes	no	yes
Newmark & Melaragno (43)			✓	all pupils in school	upper grade	all pupils in school	lower grade	>1 yr	yes, with help	yes	no	yes
Ebersole (45)			✓	"	"	"	"	>1 yr	yes	yes	no	yes
Niedermeyer & Ellis (47)	✓			112	5th/6th	112	K		yes, with help	yes	yes	no
Mohan (50)			✓	16	5th/6th	16	2nd/3rd	12 wks	no	yes	yes	"low motivation"
Olds (54)			✓		6th/adult		2nd	1 yr	yes, with help	some	yes	no
Ellson, Harris & Barber (65)		✓			paraprofessional adults	>300	1st	>1 yr	no	yes	yes	some
Ronshausen (75)		✓			"	>500	K/1st	>1 yr	no	yes	yes	some

Figure 2. Characteristics of major studies examined

CHAPTER II. COGNITIVE OUTCOMES IN EXPERIMENTS AND FIELD STUDIES

Experiments in School Settings

Experiments often introduce an artificiality into a school situation: "outsiders" intrude and help to run the experiment, many more tests than usual are given, questionnaires must be answered, and the entire flurry of activity is a short-term transitory phenomenon. It is difficult to assess the extent to which these exigencies affect the outcome of an experiment conducted in a school setting.

However, although experiments introduce an artificiality that makes generalizations from the results a hazardous procedure, they also allow for controls to be established so that, within their context, strong conclusions can be drawn. If experimental results and field study results point in the same direction, confidence in both is considerably increased.

Being a Tutor vs. Remaining in Class

Fitz-Gibbon, 1975. In one experiment that was conducted in a school, a single Learning-Tutoring Cycle* was implemented for 40 ninth grade students randomly selected from four low achieving "general math" classes in an inner-city junior high school, population approximately 90 percent Black and 10 percent Chicano. Tutees were 68 fourth grade students randomly selected from the three classes that comprised the fourth grade at an adjoining elementary school. Tutoring was conducted for a three week period and aimed at eleven objectives concerned with the addition of fractions. Immediate posttests and questionnaires were administered, followed three months later by retention tests, more questionnaires, and the Intellectual Achievement Responsibility Scale (Crandall, Katkovsky, Crandall, 1965). Both power and speed tests were used. Power tests contained five items measuring each of the eleven objectives.

---

*The Learning Tutoring Cycle is a Learning-by-Tutoring project which is a primary focus of the CSE Reports on Tutoring. See Preface.

The learning part of the LT cycle, i.e., the initial instruction, was brief: it consisted of three class periods during which the addition of fractions was explained. Methods of teaching the meaning of fractions were demonstrated and some use was made of role playing. In three of the four ninth-grade classes, tutors and non-tutors received this initial instruction together. In these three classes, the non-tutors then worked with their regular teacher on the eleven objectives throughout the next three weeks, while tutors spent the same three weeks teaching the objectives to fourth graders under the supervision of the experimenter. Non-tutors in those three classes thus constituted a competing-treatment control group: they received classroom instruction whereas the experimental group participated in supervised tutoring. The time allocated per day to supervised tutoring was either 30, 40, 60, 65 or 80 minutes per day and tutors worked with either one, two, or four tutees.

In the fourth class, tutors were simply taken out of the class both for initial instruction and for the three weeks of tutoring during which time the class continued with its ninth grade general math program. Since no effort was made to have the class work on the eleven objectives, students in this classroom formed a no-treatment or baseline control group.

Means and 95 percent confidence limits for ninth graders on the immediate posttest are displayed in Figure 3.

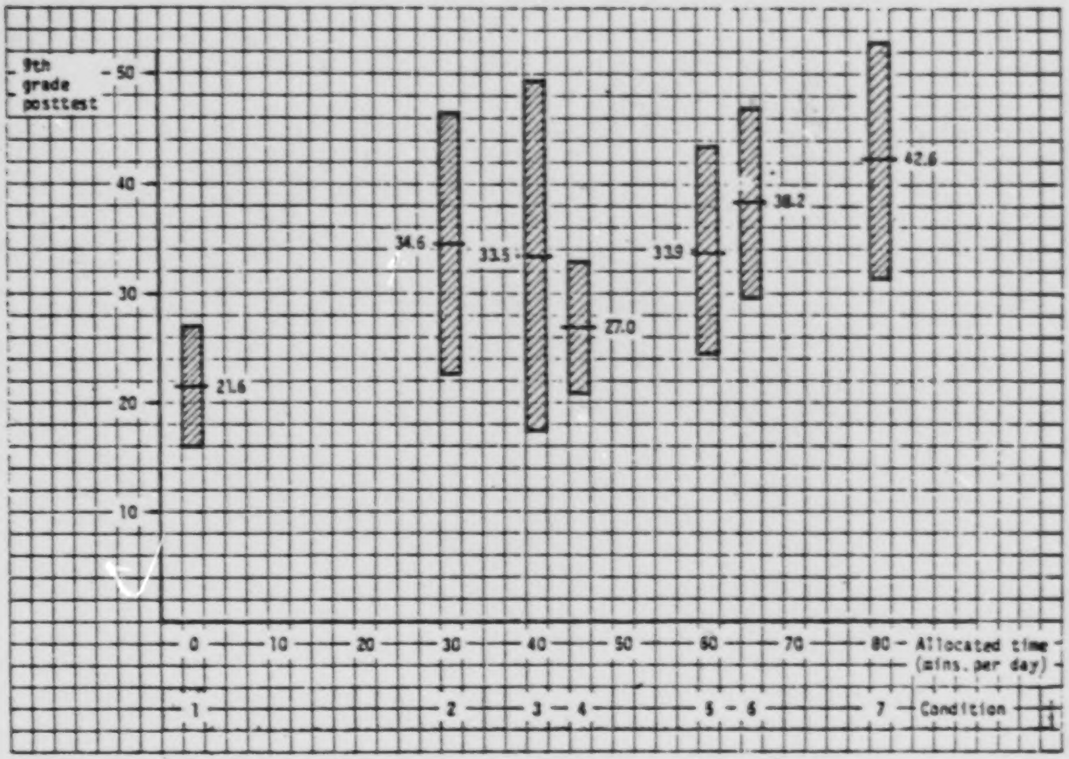

Figure 3. Ninth grade posttest results: posttest means and 95% confidence limits for ninth graders in the seven conditions.

The no-treatment control group mean ("Condition 1" in the figure) was 21.6 and the mean of the competing-treatment control group ("Condition 4") was 27.0. Experimental groups (tutors) obtained means ranging from 33.5 to 42.6. A contrast of condition 4 with the tutoring conditions was significant ($\underline{t} = 3.02$, df = 65, $p = .004$). Clearly tutors, who had spent three weeks teaching fourth graders about fractions had themselves learned more about fractions than the students practicing fractions in their math class.

Effectiveness of Low Achieving Students as Tutors

Had these low achieving students been effective in teaching the fourth graders? All fourth graders took a posttest covering the eleven objectives but employing smaller numbers than those involved on the ninth grade test. On this fourth grade test the items were randomly ordered rather than occurring in groups of five as had been the case on the ninth grade test. This randomly ordered test was called the Scrambled Fractions test. Means and 95 percent confidence limits on the means are displayed in Figure 4 for the non-tutees ("o minutes time allocated") and for tutees that had received 20, 30 or 40 minutes tutoring per day. One-way, fixed effects, analyses of variance on the posttest scores indicated significant effects [$F_{(3,90)} = 6.39$, $p < .001$].

Tests three months after the brief tutoring experiment showed tutors still with a significant advantage over non-tutors. Non-tutors, in fact, had sunk back to a mean score not significantly higher than their pretest score. Among fourth graders there had been a considerable amount of forgetting, but tutees still showed significantly higher average scores than non-tutees [$F_{(3,76)} = 4.79$, $p < .01$].

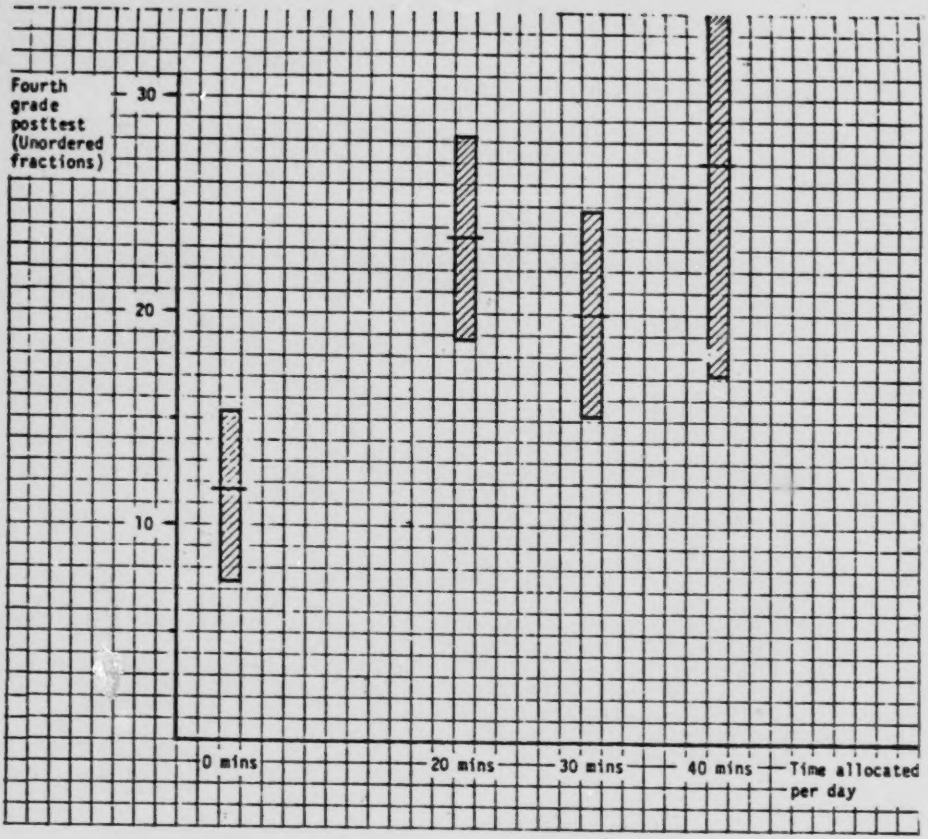

Figure 4. Fourth grade posttest results: Posttest means and 95% confidence limits for fourth graders allocated zero, 20, 30, or 40 minutes of tutoring daily.

Ninth graders were administered the Scrambled Fractions test at posttest time in addition to the ninth grade posttest. This permitted a direct comparison of ninth and fourth graders who had and had not participated in the tutoring project. The distribution of scores on the Scrambled Fractions test is shown in Figure 5. In three weeks of tutoring the ninth graders had moved the fourth grade score distribution for tutees close to that of the ninth grade control group students.

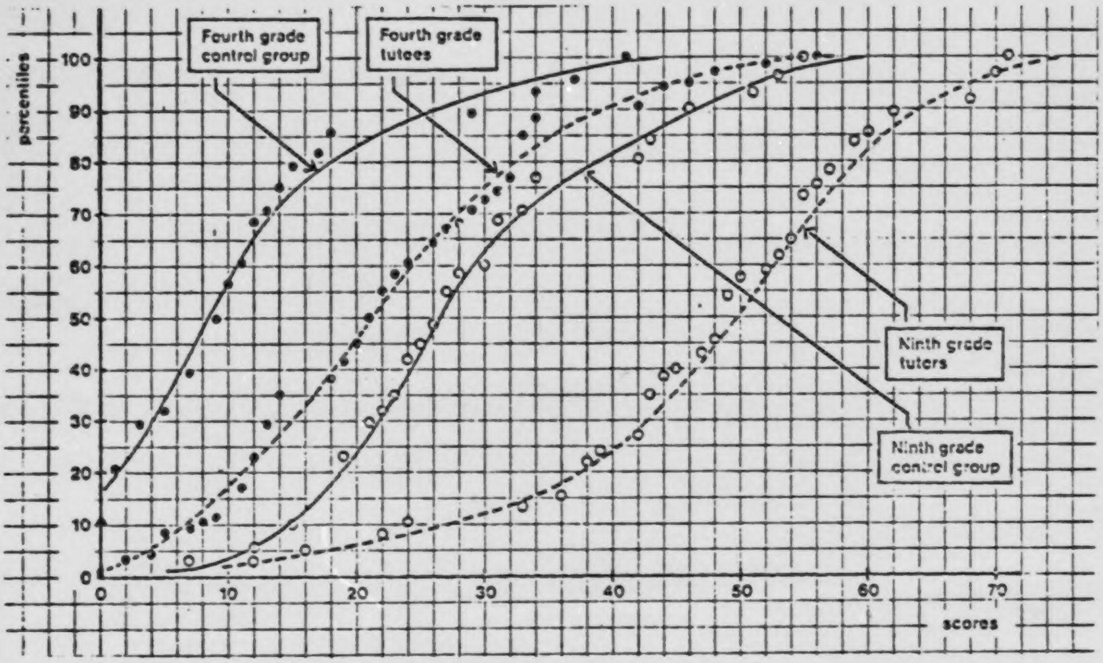

Figure 5. Scores of experimental and control groups of the Scrambled Fractions posttest.

Many practitioners with experience of running tutoring programs in the schools maintain that less able students make better tutors, perhaps because they are patient and tolerant of repetition. To examine the relative effectiveness of the tutors in this study, tutee residual gains were computed by regression of posttest scores on three variables: amount of tutoring actually received (in minutes), tutee ability, and tutee's "effort" characteristics. ("Effort" was a rating by the student's teacher of the student's usual level of academic effort in the regular classroom.) For each tutor, the average residual gain of his or her tutees was then computed. What kinds of tutors produced the highest achievement in tutees, once time, ability and effort were controlled? As a first indication it can be noted that tutees' average residual gains were correlated negatively with tutor ability (-.09) pretest

(-.14) posttest (-.15) and retention test (-.28), although only the latter correlation was significant at the .05 level. There was a significant negative correlation between average tutee residual gains and tutors' effort ratings (r = -.30, p = .03, n = 39). These trends indicated that low effort, low achieving, students were slightly more effective as teachers than were high effort, high achieving, students. The tendency was not a strong one but was sufficient to refute any suggestion of a tendency in the opposite direction for this group of students, a group, it must be remembered, of generally low achievement. These findings were not inconsistent with the observations of many practitioners: poor students are frequently relatively effective tutors.

How can persons less competent in the subject be as effective or more effective when acting as tutors? In considering answers it must first be recalled that tutoring is not teaching. Tutoring means explaining a limited, usually prescribed, set of concepts and very often it is performed following training. Had tutors the responsibility for selecting curriculum, analyzing and organizing it into clear objectives, then one might expect subject matter competence to become important if not crucial to success. But in the context of clearly delimited instruction, and with training, low achievers can make effective tutors.

(An indication of one way in which child tutors may be effective is obtained from some reports by Stallings and Kaskowitz (1974) and Soar (1972). They presented evidence from large-scale analyses that open-ended questions, i.e., questions high in the Bloom taxonomy (Bloom, 1956), are negatively related to achievement. Direct questions at lower levels of the Bloom taxonomy facilitate acquisition of knowledge by students. Lower achieving or less competent tutors may be likely to ask simple direct questions rather than indirect, abstract, reasoning-type questions.)

In summary, the results of this experiment conducted in a school setting indicated that after ninth grade students had received initial instruction together, acting as a tutor produced significantly higher achievement than did classroom instruction. Fourth grade tutees made significant gains as a result of tutoring and both tutors and non-tutors still showed significant advantages over control groups three months after the three week experiment. There was a slight tendency for low-effort students to have been more effective as tutors than students whose teachers rated them higher in usual academic effort.

Being a Tutor vs. Studying Alone

Allen and Feldman (1973). The hypothesis that low-achieving children will learn more when placed in the role of teacher than when spending the same amount of time studying alone was tested by these researchers.

Tutors were 10 low-achieving fifth graders and tutees were 10 third graders (small numbers, but it was a repeated measures design). All children were paid to participate in the experiment during the summer vacation. Tutors and tutees were sex-matched.

The tasks were 10 different lessons dealing with science, language, and reading stories. The experiment ran a fortnight with tutoring taking place each weekday. Each day the fifth graders studied the lesson for eight minutes and then either continued to study alone for the next 20 minutes or spent the same amount of time tutoring. The "study alone" and "tutoring" conditions were assigned on alternate days to each fifth grader. Thus each day half the subjects were tutoring and half studying alone. The third graders were assigned on alternate days to the tutee role or to study alone. An objective test on each lesson was given at the end of each session.

Test scores were converted to z-scores and the results are shown in Figures 6 and 7 , reproduced from the article. (The graphs are based on only seven subjects due to experimental mortality: one tutor did not attend all sessions and two did not follow instructions.)

Analysis of variance for repeated measures indicated that for the younger children there was an interaction between condition and session significant at the .10 level but no significant differences in orthogonal comparisons.

For the older children (tutors) the interaction between condition (tutor or study alone) and session (occasion) was significant at the level of $p < .08$, and there were significant differences at the .05 level in favor of tutoring at sessions 4 and 5. Thus, as children became used to tutoring, their performance improved relative to that obtained under the study alone condition.

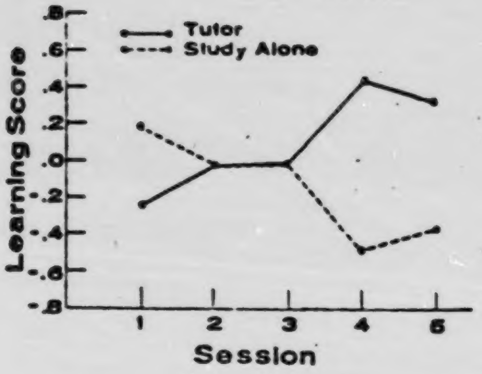

Figure 6. Allen and Feldman results for tutors: Standardized scores for retention of material when studying alone versus tutoring (fifth graders)

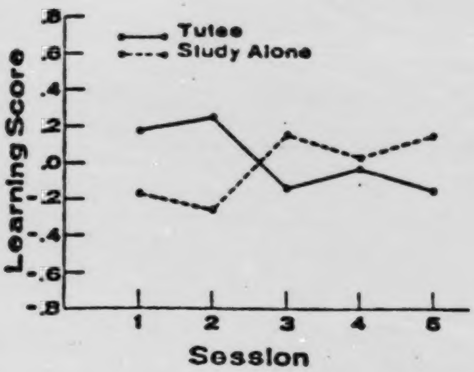

Figure 7. Allen and Feldman results for tutees: Standardized scores for retention of material when being taught versus studying alone (third graders)

The authors concluded that the results supported the hypothesis that low-achieving children will learn more when assigned a tutor role than when spending the same time studying alone.

Possible Processes Involved in Promoting Learning in Tutors

The two experiments just described implied that acting as a tutor led to more learning than studying alone (Allen & Feldman, 1973) or working in class (Fitz-Gibbon, 1975). What processes may account for such effects? Allen and Feldman posed the following problem:

> Of considerable theoretical importance is the determination of the locus of the enhanced performance due to tutoring. As one possibility, perhaps the mere expectancy of teaching someone else is sufficient to produce increased learning--thus, the tutoring situation itself might be of only secondary importance. The expectation of tutoring could have resulted in the child's exerting greater effort and working harder on the material while preparing the lesson (p. 4).

In other words, does the act of tutoring, in itself, promote learning or is the beneficial impact of the tutor role accounted for solely by the initial motivation to learn the work in order to teach it?

Allen and Feldman indicated that this is a question of "theoretical" importance. Not so; it has immediate policy-related implications: if the motivation to learn rather than the process of tutoring is the major source of influence on cognitive gains, then presumably other kinds of motivation could replace tutoring and be equally effective. Motivation might be induced by attractive and meaningful instructional materials, livelier teachers, or material rewards for example. If, however, there is something about the activity of tutoring that produces cognitive gains, then one must insist that this activity occur, and reject as unpromising a search for other forms of motivation to produce equivalent gains.

If the motivating effect of knowing one is to tutor is the source of a major part of the cognitive gains experienced by tutors, then an assessment of tutors and non-tutors after initial instruction but prior to tutoring should show tutors already ahead of non tutors. Both Allen & Feldman and Fitz-Gibbon made such measurements. In the Allen and Feldman study, tutors and non-tutors studied the lessons on their own for eight minutes after which initial preparation time each made a tape recording of what they had learned. These recordings were analyzed in an attempt to detect better free recall and organization among those who were to be tutors. No significant differences were found in content analysis or in the total number of words produced. Fitz-Gibbon's results were similarly negative: no significant differences between students who knew they were about to tutor and students who were not to be tutors on a test given immediately following the three days of initial instruction. It was during the three weeks of actual tutoring that tutors moved ahead to a significant advantage over non-tutors.

These two findings suggest that the source of the effect of tutoring on the tutor resides, primarily at least, in the act of tutoring rather than in the initial motivation deriving from the expectation of tutoring. It is of course possible that this expectation could be providing a weak effect that the experiments were not powerful enough to discover. It is also possible that the arousal effect of knowing one will tutor could increase with experience. Nevertheless, present scant evidence suggests that the act of tutoring is itself the major locus of the cognitive gains shown by tutors.

What is it about the tutoring process that produces cognitive gains in the tutor? One feature that distinguishes the tutor's activity from that of students in other roles is the verbal activity involved in tutoring: the posing of questions, the generation of explanations. This verbal activity might be the source of the superior learning exhibited by tutors, a proposition for which some supporting evidence will be cited shortly. It is possible that by verbalizing, one is alerted to errors in one's own reasoning or to aspects of a problem that would otherwise pass unnoticed. Many persons report the experience of struggling with a problem, verbalizing it by asking someone's help, and then immediately seeing the solution without help; posing the question sparked the solution. Conversely, one may believe one knows some concepts thoroughly. Someone asks for an explanation and, after a confident start explaining the work, one notices a small non sequitur which raises more issues and leads to a realization of a new aspect of the problem, one yet to be understood.

Merz (1969) examined different conditions for the solution of the Progressive Matrices problems. (The Standard Progressive Matrices [Raven, 1958] is a test of reasoning employing items that consist of a matrix of patterns with one cell missing. The examinee must choose the design that completes the pattern inherent in the matrix. The test has been widely used as a "culture-fair" test of reasoning.) Subjects instructed to reason out their answers aloud obtained significantly higher scores than subjects working the test silently. The whole body of literature on overt responding lends support to the idea that the verbal-activity involved in tutoring might be an important source of the gains.

On the other hand, the superior learning might be due to the sustained motivation provided by tutoring. Initial, short term motivation might not differ between tutoring and non-tutoring conditions but tutoring might pall less rapidly than other practice situations. The sustained motivation would be evidenced by more time-on-task behavior engaged in by the tutor during tutoring sessions, than by non-tutors who were in classes or studying alone. No measures of this important, observable variable have been made. The time-on-task hypothesis must remain as an untested competing hypothesis for the verbal-behavior hypothesis. (The question, of course, is not necessarily an either/or question, but one of relative magnitudes of the contributions made by time-on-task effects, by verbal activity effects, and by initial motivation to learn.)

Being a Tutee vs. Being in a Class

The preceding pages have concentrated largely upon the cognitive benefits of being a tutor. Fitz-Gibbon's data indicated benefits to tutees (children tutored by ninth graders) in contrast to a baseline, no treatment control group, but Allen & Feldman's data showed the younger children did as well studying alone as receiving tutoring from fifth graders. An experiment reported by Bausell, Moody and Walzl (1972) compared achievement of tutees after tutoring by adults with the achievement of a control group that received an equal amount of instructional time in a classroom.

Bausell, Moody and Walzl (1972). There were 20 experimental teachers: ten under-graduate sophomores and ten student teachers. Each was assigned to teach six fourth and fifth grade pupils. Stratified random sampling was used to select groups consisting of two high, two medium, and two low ability pupils in order to ensure that the teachers taught similar groups of six pupils. In each group, three pupils (one high, one medium, and one low) were randomly assigned to be taught individually by the teacher and three were taught in a class by the same teacher. During the classroom teaching, teachers did not know that some pupils in the class were experimental subjects. The task was an elementary unit on exponentials and a 16-item posttest was administered immediately following the 30 minutes of instruction.

Analysis of covariance showed a significant effect in favor of students who had received tutoring. On the 16-item posttest this effect amounted only to about one item. However, this small effect occurred consistently and represented an effect gained after just 30 minutes of instruction. There were no significant differences between the results obtained by the undergraduates with no teaching experience and the student teachers who had had teacher training and some practice in classroom teaching.

All experimental teachers tutored their three pupils before they taught the class in which the three other pupils were embedded. This sequence should presumably have worked against the hypothesis that tutoring was the more effective instructional method since teachers would have had practice with the unit by the time they instructed the class.

Bausell et al. commented that "classroom instruction is obviously more efficient in terms of total learning produced per given unit of instructional time." Such a statement is only correct if we view the tutor's time as a cost to the system and do not include tutor learning as a benefit to the system. This kind of assessment of costs and benefits is appropriate when tutors are paid adults and the goals are limited to tutee gains. In the Learning-by-Tutoring projects, however, the tutor's time is not a cost since the tutors themselves are benefitting.

One of the claims made in favor of the use of peer-group or cross-age tutors is that, because of language differences, tutees may understand the instruction provided by such tutors better than they understand their teachers' instruction. Tutors speak a language, it is hypothesized, that is more closely related to that of the tutees than is the language of the teacher.

The influence of a few unfamiliar words on comprehension was demonstrated in an experiment by Marks, Doctorow and Wittrock (1975). They varied the "meaningfulness" of 20% of the words in a story. Story comprehension and retention were almost halved by the presence of 20% unfamiliar words. A beginning teacher often makes the mistake of using many unfamiliar words. Gradually the teachers who are observant of student reactions, learning, and comprehension, learn to select words with care from a limited vocabulary (at least in situations where explanation, not vocabulary building, is the aim). When teachers are not only considerably older and more educated than their students but also of different cultural backgrounds, this monitoring of their own speech must be severe. A tutor, on the other hand, has to exert little or no such effort because his or her vocabulary is not much different from that of the tutee. The Marks, Doctorow and Wittrock experiment suggests that this language match might be of substantial importance to tutee comprehension.

Misleading Claims

An experiment titled "On teaching disadvantaged preschoolers to read: A successful experiment" (Hamblin and Hamblin, 1972) purports to show that peer tutoring and reinforcements for reading are more effective than adult tutoring and reinforcement for attendance. The authors claim that their findings are "particularly significant because tutoring by an adult and privileges for attendance is what has been thought of traditionally as the ideal learning environment."

On examination, it turns out that "adult" tutoring was provided by job corps teenagers. Since it is recommended that the LTC be used with secondary level students, it is important to examine the results of an experiment which claims they are less effective than peer-tutors (preschoolers), especially since this experiment was conducted in a "disadvantaged" population.

The researchers investigated the independent and combined effects of two factors, token reinforcement and tutoring, on the reading progress of disadvantaged, inner city preschoolers, 26 white and 6 black. The two levels of the tutoring factors were defined by the characteristics of tutors. Tutors were either those preschoolers in the classroom who had learned quickly or one of five teenage job corps workers. The two levels of the token reinforcement factor consisted of tokens for either attendance or reading.

The 32 preschoolers were randomly divided into four groups to fit the cells of a 2 x 2 factorial design with the factors being the kind of behavior reinforced and the kind of tutor.

Results showed significant main effects for each factor in favor of preschool rather than teenage tutors and reinforcement for reading rather than for attendance. The results which were the most task-relevant are displayed in Figure 8. (As already noted, the label "adult tutoring" in this figure refers to tutoring by the job corps teenagers and "peer tutoring" refers to tutoring by preschoolers.)

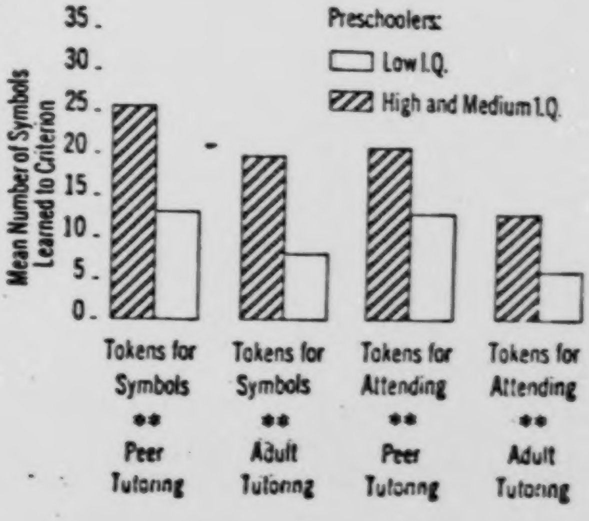

Figure 8. Hamblin and Hamblin results: The relationship of symbols learned to tokens and tutoring by IQ levels.

At first glance this looks like an experiment which casts grave doubt on the ability of teenage tutors to produce gains in preschoolers, but this interpretation cannot be accepted for several reasons. First, it is necessary to take a closer look at the nature of the treatment conditions. The authors stated:

> During the eight weeks of the experiment all of the children in two of the groups in the "adult-tutoring" condition were tutored by the Job Corps teenagers. In the other two groups, the faster children were again tutored, but the slower ones were tutored by the faster children in their group. This latter [situation] is referred to as the peer tutoring condition.

Thus, the peer-tutoring condition was only peer tutoring for the slow students. "Fast" students in this condition not only received daily teenage ("adult") tutoring but also spent an extra session per day (or more perhaps if one fast student tutored several slow children) working on the tutoring. Time-on-task was thus at least doubled for some "fast" students in "peer tutoring groups."

Furthermore, the "tutoring" consisted of sitting beside a child who was working on a Language Master, a machine which sounds out symbols when the child puts a card into it. The "tutor" had only to ensure the correct use of the machine: that the child sounded out the symbol before putting the card in the machine to see if he or she was correct. The process involves at best some monitoring and reinforcement, but it does not require the teaching of any content. It could be that the presence of teenagers was more distracting to the preschoolers than the presence of children of their own age who had just been working on the same machine, and it was this distraction that caused the lower scores. However, this effect cannot be taken to imply that the preschooler could _teach_ better than the teenager, had there been any teaching required.

Caution re Standardized Tests

An experiment designed to provide data on the efficacy of tutoring employed randomly constituted control groups of tutors and tutees and was run for eight weeks (Snapp, Oakland and Williams, 1972). Fifth and sixth grade students tutored first, second and third grade students in a disadvantaged population. Tutoring took place 20 minutes per day four days per week and was conducted before school. A "Word Recognition Test" developed from the materials used in the study assessed tutee outcomes in comparison with the randomly constituted control group and indicated significant advantage to tutees ($F_{1,58} = 8.09$, $p < .01$). A standardized reading achievement test (the Metropolitan Achievement Test), however, showed no effects of the tutoring on tutees' reading scores. Tutors were assessed only on the standardized tests and showed no significant advantage over the randomly constituted control group.

These results are entirely consistent with the position that students learn the specific content of tutoring lessons but that the general content of standardized achievement tests is not likely to detect the effects of these specific learning gains with any sensitivity. One might expect eventual gains after perhaps a year or two of consistent progress, but certainly not after eight weeks.

Field Studies with Secondary School Tutors

In continuing to examine cognitive outcomes of tutoring, we turn now to programs implemented in natural settings and continued over sufficiently long periods of time for novelty effects to wear off. Examined first are projects that involved secondary school students as tutors. Tutoring programs involving use of upper grade elementary school students are then considered.

A good proportion of the growing interest in tutoring in the last decade or two can be traced back to the work of Ronald and Peggy Lippitt at the University of Michigan, work which started in the early 1960s and continues today. Their emphasis has been on the affective impact of the helping relationship that is inherent in tutoring. Their ideas influenced two of the largest tutoring projects: Youth Tutoring Youth and the Ontario-Montclair project. In Youth Tutoring Youth programs, disadvantaged secondary school students are paid to tutor elementary school children either during the summer or in after-school hours during the school year. Cloward (1967) presented a thorough evaluation of the New York Homework Helpers Program that ran on this model.

A Youth Tutoring Youth Project

Cloward's data were based on a 26-week program during the 1963-1964 year. Eleven tutorial centers were established in New York City elementary schools, each under the supervision of a licensed teacher. Tenth and eleventh grade students were paid $11.00 per week for six hours of tutoring in Reading and Language Arts and two hours of inservice training. The "central issue" was described as "whether students who have not completed high school can serve effectively as tutors for low-achieving public elementary school pupils."

Tutors were recruited through a publicity campaign which included flyers to local academic and vocational high schools. Most of the 240 tutors came from the academic high schools of the area in which the tutoring centers were located. Sixty-one percent of the tutors were white, and the remainder consisted of approximately equal numbers of black and Puerto Rican students with a few (2%) Oriental. Sixty-three percent of the tutors were female. There were similar numbers of tenth and eleventh grade students (52% were tenth grade), and the mean reading level was 10.0. No students were employed as tutors if they were reading more than three years below grade level. School guidance counselors had been asked to screen applicants who were "emotionally disturbed" or in "serious academic jeopardy."

Tutees, who were fourth and fifth grade pupils, were randomly selected from a pool of students who were reading below grade level and whose parents responded to an invitation to apply for the tutorial program. However, students with problems such as poor behavior, truancy, or retarded mental development were screened out by counselors before random selection was used to form an experimental group of 410 pupils and a control group of 185 pupils. Attrition, due to the failure to attend of 54 experimental pupils (13% of the experimental group) and 28 control group pupils (15%--no data were available for the 24, and 4 were inadvertently used to replace non-reporting tutees) reduced the numbers to 356 tutees and 157 non-tutees (the control group). These equivalent groups were composed roughly of 10% white, 30% black, 57% Puerto Rican, and 3% Oriental, and of even numbers of male and female.

Eight afternoons were devoted to the preservice training of tutors. "These preservice training experiences focused on (a) the goals of the tutorial program, the organization of the program, and the duties of a tutor, and (b) characteristics of the youngsters they would encounter, the scholastic levels at which these pupils would be operating, and the kinds of activity in which tutors might engage their pupils during the first few meetings."

Throughout the duration of the program, tutors received in-service training two hours per week. Early in-service meetings tended to concentrate on the specifics of the elementary curriculum and the problem of developing good tutor-pupil relationships. Later meetings were focused on the tutor's anecdotal reports, which were used as case studies to stimulate discussion of problems encountered in the tutorial situation.

Tutees were assigned to tutor in one of three ways: matched by sex and ethnicity, matched by sex only, not matched in sex or ethnicity. Tutoring sessions were held from 3:00 to 5:00 p.m. on Tuesdays, Wednesdays, and Thursdays. Each tutoring pair met either once or twice a week. Thus there were two "exposure levels" for tutees. It is not stated, but it appears from the numbers that tutors sometimes were assigned two tutees to tutor on separate days. Tutoring was conducted in classrooms with up to four pairs working in a room, a very non-crowded condition. The pattern of a typical tutorial session became to spend about one hour on reading and homework and the remaining time on games, refreshments, roll taking, "and other non-tutorial activities." After the session, tutors escorted their tutees home. Tutees were thus receiving either one or two hours a week academic help and an equal amount of non-academic attention. As mentioned, tutors met for two hours weekly with the supervising teachers in addition to tutoring.

Reading achievement of tutees was pre- and posttested by the New York Tests of Growth in Reading, Level C, form 1 (revised), with the results reported in Table 1.

TABLE 1

Cloward's Results for Tutees

Sample	*N*	Pretest mean	Mean gain
2 tutoring sessions a week	100	21.33	+6.20
1 tutoring session a week	73	22.89	+5.33
Controls (no tutoring)	79	21.98	+3.48

Covariance Analysis

Source	Before ss	After	*df*	*MS*	*F*	Required .05
Between	331.87	332.31	2	166.15	3.68	3.04
Within	11,603.84	11,022.59	244	45.17		
Total	11,935.72	11,354.90	246			

Factors used as covariates were sex, ethnicity, school grade, access to school programs in reading remediation, and pre-study reading level. Applications of the Duncan Range Statistic indicated a significant difference in means occurred only between the two sessions-a-week tutees and the no-tutoring controls. The author reported unsuccessful attempts to detect the relationships between the tutorial experience and school marks, behavior, attitudes, and aspirations.

A significant effect, however, was found for the tutor-pupil matching variable and an almost significant ethnicity-matching interaction. Examination of means indicated that black tutees matched to their tutors for both sex and ethnicity achieved significantly higher reading scores than control group black students.

A random selection process had been used to form a true equivalent control group of 72 students who were told they would be hired as tutors the following year provided they participated in the testing program. Using analysis of covariance, the corrected mean changes from pre- to posttest shown in Table 2 were obtained.

The tutors appeared to have made consistent gains over the control group on the reading subtests. The statistical significance level for the total score advantage of tutors was .001. What was the educational significance of the gain?

The author reported a translation of the raw score gains over the seven months between pre- and posttest to grade equivalents: these showed a gain of 3.4 years for the experimental group and 1.7 years for the control group.

However, grade-equivalent scores are difficult to interpret unless the exact method of their derivation is known. From Table 2 it can be seen that the tutors gained 13.2 on a standard score scale as opposed to a 7.6 gain by controls.

TABLE 2

Cloward's Results for Tutors:
Summary of Covariance Analysis
of Change from Prestudy to Poststudy,
Iowa Silent Reading Tests (Standard Scores)
for Experimental and Control Subjects

Test	Corrected Mean Change[a]		
	Experiments	Controls	F
Rate	-4.9	-6.4	0.12
Comprehension	9.0	1.4	7.76*
Directed Reading	24.8	15.0	9.20**
Poetry	12.0	7.9	2.57
Word Meaning	12.6	11.4	0.23
Sentences	9.9	10.3	0.05
Paragraphs	15.3	10.1	4.98*
Use of Index	18.7	13.8	3.67
Key Words	13.3	9.6	2.39
Total (Median Score)	13.2	7.6	14.14***

[a] Covariates used were prestudy reading level, Quick-Word score, sex, and school grade.

*Significant at .05

**Significant at .01

***Significant at .001

The attrition rate was 37.4% among tutors and 20.8% among non-tutors. Twelve percent of the tutors failed to report for the tutoring job and 23% resigned during the program. It was stated that most of those who resigned had class changes at the semester which prevented their tutoring. Whether the tutors who remained in the program to posttest time were high aspiring, determined students who would have made considerable progress anyway cannot be known.

The author made the following summary statements:

> It is clear from the findings of these studies that to be effective, tutors do not need 12 years of formal education and extensive training in reading pedagogy. Nor need they be highly successful in their own school work. The average high school student can learn to be an effective tutor.
>
> Clearly, the major impact of the tutorial experience was on the tutors themselves. This finding has implications both for education and for youth employment. Tutorial programs not only can provide older youth in a low income area with gainful employment but can serve to upgrade their academic skills as well.
>
> Indeed, the high reading gains made by tutors many of whom were reading far below grade levels at the beginning of the study, raises the intriguing question of whether high school dropouts might be successfully employed as tutors, not just to help underachieving elementary school pupils, but to improve their own academic skills.
>
> Attempts to remedy the dropout's educational deficiencies by placing them in pre-employment training programs have not been notably successful. Having experienced failure and humiliation in the classroom and being alienated from school, these youngsters tend to rebel against learning situations in which they are cast in the role of a student. Assigning tutorial roles to such adolescents might help to make learning enjoyable and profitable for them (p. 24, emphasis added).

(Cloward's finding that black tutees matched to their tutors for both sex and ethnicity achieved significantly higher reading scores than control group black students was of interest. To what might such an effect be attributed? Several writers have suggested that student tutors can be more effective than adults because the tutee is more willing to ask questions, more relaxed in his or her relationship, less defensive with a fellow student than with an adult. The authority figure which an adult represents may inhibit learning in some students. Perhaps black tutees displayed a similar defensiveness when assigned to tutors of other ethnic groups.)

The relevance for the present report of the program described by Cloward is that it was a large scale field implementation of a secondary-to-elementary tutoring program. More of the treatment appeared to make more of a difference for tutees, indicating that secondary school students can assist the learning process for elementary school children. Benefits in reading achievement seemed to accrue to tutors, but of course they spent extra time on reading each week and tutoring was not compared with any alternative use of the tutors' time. However, the tutors' school work did not suffer as a result of their participation as paid tutors. The treatment was not very intensive: as little as one hour of instruction per week could hardly be expected to build up much momentum. In the paragraphs quoted above, Cloward clearly stated several of the hypotheses which encourage the implementation of the Learning Tutoring Cycle.

The Youth-Tutoring-Youth project is run during out-of-school hours and tutors are paid. The YTY program is one for which successful tutors from LTC projects might be made eligible. Work as a tutor during the school year followed by employment as a tutor might be a synergistic arrangement.

Three Cross-School Tutoring Projects

We turn now to another project strongly influenced by the work of the Lippitts: the Ontario-Montclair project. This California project won an NIE pacesetter award and a dissemination grant under Title III ESEA. An evaluation of the project was performed by an external consulting firm, but their report was inadequate. Tutors were clearly volunteers and there was no mention made of how a control group was formed. Differences between experimental and control groups were reported by significance levels rather than by actual data. However, the method of implementation of these projects is of interest, and personnel involved are convinced that students benefit from improved self-concepts and learning skills. Many instances of dramatically improved attendance are cited in connection with these cross-age tutoring projects..

In the typical Ontario-Montclair project, participation as a tutor is voluntary, but project directors have generally made an effort to recruit both low-achieving students and high status students such as football stars. Heavy emphasis is placed on understanding the tutees' problems and establishing warm relationships with the tutees. Tutors are encouraged to prepare their own games and materials for use in tutoring and typically spend their Monday class period on such preparation activities. They tutor on Tuesday, Wednesday and Thursday, then devote Friday's class period to "feedback"--discussing problems encountered during the tutoring sessions.

While tutors form an entire class at the secondary school, tutees are usually pulled from several classrooms and tutoring takes place at various locations around the elementary school, such as cafeterias, hallways, lawns. The tutee's teacher often assigns work; otherwise the tutor decides what is taught. This very flexible kind of program would be difficult to evaluate for cognitive impact since the tutoring is highly individualized. The main emphasis is firmly on the helping relationship.

Programs similar to the Ontario-Montclair project have been reported from Madison, Wisconsin and Paducah, Kentucky (Klaus, 1975, pp. 51-60). In both these programs secondary students (primarily eighth or ninth grade students) can elect to be tutors at elementary schools. Again, the arrangements made are flexible: tutors are assigned individually to various receiving teachers and how they work depends very much upon the wishes of the receiving teacher. Supervision is a major problem in such programs. The receiving teachers have their regular classes to supervise, and the coordinator cannot be in a dozen classrooms all at the same time. The "coordinator" is frequently a guidance counselor with other responsibilities.

In Paducah, although the primary concern of project personnel was to enhance the social attitudes of participants, assessments of achievement were made by the Board of Education during the 1973-74 school year. For the 320 tutors involved from two junior high schools, gains during the year were half a month less "than expected" in reading skill but four months more "than expected" in mathematics. Interpretations of such data are difficult, and no correction can be made for the initial self-selection process of the volunteers. It would be interesting to know if tutors worked in math more than in reading, but data collected in the framework of a good design is needed for firm conclusions. However, the fact that tutees also gained more than expected in math but not in reading fits in with the supposition that math tutoring was occurring and producing noticeable effects. (A narrative account of the Paducah project can be found as an appendix to CSE Report on Tutoring #3.)

A Within-Secondary-School LbT Project

In contrast to the highly flexible cross-school Teacher Aide projects, in which tutors are volunteers and are assigned one or two at a time to various elementary school classrooms, a master's degree thesis prepared by a junior high school teacher illustrates a Learning-by-Tutoring project (Hoffmeister, 1973):

In an inner-city junior high school (approximately 90 percent Black), two ninth grade classes tutored two seventh grade classes in reading throughout the second half of a school year. The ninth grade tutors received pay. This was an elective course, and there were almost twice as many females as males. Tutors were, on average, two years below grade level in reading--which, however, placed them slightly above the average for the school's ninth grade.

The seventh grade tutees were selected because they were the most deficient in reading skills, reading on average at a third grade level.

The entire first semester was devoted to tutor training (12 weeks) and practice tutoring (6 weeks). A serious task-orientation was established from the beginning. Tutors learned, during training, to assess reading grade levels, check skills (such as blending, the recognition of short-vowels, and consonant combinations) and to test comprehension. Students assembled materials such as diagnostic tests and became familiar with all the reading resource materials. They learned to make daily lesson plans and made extensive use of role playing to practice the skills being acquired. Following the 12 weeks of training, tutors undertook practice tutoring, for no pay, for 6 weeks. The ninth graders were paired with seventh graders and tutored for one class period per day, four days a week. At the beginning of the second semester the tutoring arrangement continued with the additional incentive of pay for tutors ($1.67 an hour).

Hoffmeister wrote:

> Definite rules were established for receiving pay each day. They were as follows:
>
> 1. Tutors must be in the class for the entire period.
> 2. First period tutors (beginning of the school day) cannot be more than ten minutes late to class if they want to be paid for that day.
> 3. Tutors must have a written daily lesson plan at the beginning of each period.
> 4. Tutors must carry out their lesson plan quietly, so the classroom can retain a working atmosphere.
> 5. Tutors must co-operate respectfully with the supervisors.
>
> The paycheck was an exciting part of the program to these young people, most of whom had never earned this much money. It was also a sustaining factor in the program. When things got more routine in April and May, as any project will with time, the paycheck was a big incentive to keep up daily work. Very occasionally the statement of "No work - no pay" had to be made to a lagging or unprepared tutor. And even less frequently this writer found it necessary to delete their attendance mark from the daily time chart which meant they would not receive pay for that day.

> Another very important part of the peer-tutoring program was the Monday Seminar. This first day of each school week was not used for tutoring, but was set aside for a working seminar session with the tutors. They met as a group in Room 54 with the writer, while the tutees met as a group with Miss Sussman. This day was a day for continuing tutor-training, an in-service workshop so to speak. Lesson plans which had been turned in the previous Friday and checked over the weekend were discussed individually with each tutor, common problems or errors in planning were discussed with the body of tutors, new material was introduced and distributed, tutors shared problems with each other for encouragement and advice, lesson plans for the coming week could be outlined, and materials for them gathered.
>
> This seminar became a vital part of the program. It was the day when all of the tutors could see each other and share the experience of which they were a part. For the supervisor it was also a vital day, because it was the only day she could see them all together, maintain a teaching relationship with them, and listen to their needs and desires. (pages 20-21)

Two forms of the Gates-MacGinitie Reading Test, Survey E were administered, one as a pretest and one as a posttest. Tutees were evaluated on the Comprehensive Test of Basic Skills. Hoffmeister presented grade equivalent scores at pretest and posttest for all students. While maintaining strong reservation about the use of standardized tests and grade equivalent scores, the results are summarized in Table 3.

TABLE 3. Hoffmeister's Results for Tutors and Tutees

		Total Reading		Comprehension		Vocabulary	
		pre	post	pre	post	pre	post
TUTEES	mean scores	3.3	4.2	3.4	3.8	3.51	4.75
(n=28)	standard deviations	(1.45)	(1.42)	(1.39)	(1.23)	(1.56)	(1.77)
	correlated t-values	t = 6.37**		t = 2.12*		t = 4.72**	
TUTORS	mean scores	7.37	8.34	6.68	7.43	8.10	9.24
(n=27)[1]	standard deviations	(1.99)	(2.07)	(2.40)	(2.22)	(2.27)	(2.34)
	correlated t-values	t = 3.52**		t = 2.54**		t = 2.21*	

* $p < .05$

** $p < .01$

[1] Three scores were eliminated from the tutor data: one was a student who did not earn any pay because of erratic attendance. She missed the posttest, and given it as a make-up, took it unwillingly and showed a 3.6 grade equivalent drop. Two other scores were eliminated because they were at the ceiling of the test at pretest.

Both tutors and tutees made significant gains of about a year in grade equivalent scores. Since tutees were selected on the basis of low scores, some regression effects had to be expected,but these scores would normally be considered "successful" in a Title I program. Hoffmeister noted that ten of the fifteen poorest readers gained 1.2 years or more and all but two made at least month-for-month gains.

A Project that Nearly Failed

Projects which fail are not usually written up for publication, which is unfortunate since practitioners need to know what fails as well as what succeeds. Dreyer (1973) reported a project which nearly failed but was saved. Since it was a cross-age tutoring project using Title I funds and employing secondary students to tutor in elementary school classrooms, it is a study relevant to the LTC proposal.

Tutors were seventh graders who were having reading difficulties. They tutored first grade children who were likewise having difficulties. Tutoring lasted 15 to 30 minutes and took place in the first-grade classrooms. First-grade teachers, meanwhile, were busy with the remainder of the class. This was identified as one of the problems at the beginning of the program: tutors were ignored. They were given no feedback or reinforcements.

The training of tutors, or rather the lack of it, was reported to have been another problem that threatened the continuance of the project.

> Some [receiving teachers] had to spend so much time instructing tutors that they could have saved time by doing the tutoring themselves. Others gave assignments so vague that the tutor did not know what he was to do. In some cases the two or three tutors assigned to each first-grade classroom were definite distractors (p. 810).

A reading resource teacher and a consultant then set about developing clear, very specific procedures for tutors, based on discussions with teachers. Much of the article described these procedures for tutoring in reading.

The method developed by teachers, the reading specialist, and the consultant for conducting the tutoring program was one involving daily prescribed assignments. Following the teaching of some material by the receiving teacher, this teacher would then indicate appropriate "repetitive practice" for the tutors by marking an assignment sheet. Tutors picked up their assignment sheets and materials and conducted the sessions with the tutees, after which each one recorded the extent to which the tutee had completed the assignment. Cumulative graphs were made to chart progress, and the resource teacher met regularly with tutors to praise and encourage them and, in particular, to urge tutors to praise and encourage tutees.

No data were presented, but the responses of teachers once the program was working were stated to have been highly positive with regard to tutor and tutee self-concepts. Significantly greater gains in reading skills among tutors and tutees were mentioned but not documented.

One can imagine the problems at first. Unprepared, low-achieving seventh graders were sent to work with low-achieving first-grade students. They were given vague assignments and ignored while the teacher continued to work with the rest of the class, the students not having difficulties.

The intervention of the reading specialist and consultant resulted in the provision of systematic procedures to

- provide tutors with specific assignments
- provide tutors with audiovisual skills and necessary materials
- provide tutors with teaching techniques
- provide tutors with encouragement and feedback
- provide tutors and tutees with progress charts

In the proposed LTC programs these are components of the Tutor Support System which has been noted as an essential feature (Report #2).

Ad Hoc Remediation vs. Full Sequence Instruction

The most usual secondary to elementary tutoring projects have been Teacher Aide projects. Secondary students are assigned one or two per room to work with selected tutees, often for remedial purposes. This pattern is evident not only in the projects described here from available literature but also from the survey, site visits, and other contacts made by the staff. Perhaps this pattern has developed because of the influence of a few projects or because it is one that fits in well with the present structure of schools. It requires only that receiving teachers accept a few aides, help for which they are often very grateful.

The LTC proposal places greater demands for change on the schools: To receive a whole class of tutors, for example, is a different proposition from receiving one or two aides. In the Learning-Tutoring Cycle, it can be argued that both more is expected of tutors and also less is expected of them. More is expected of tutors in that they are given responsibility for actual teaching, not just asked to listen to someone read or to grade papers or perform other chores that are assigned to Teacher Aides. When tutors in Teacher Aide programs are asked to teach, they must be prepared to teach whatever part of the curriculum the receiving teacher identifies as important for the tutee. Frequently, this is a part of the curriculum with which the tutee has already developed problems. Identifying misconceptions and re-educating a tutee, trying to build on what the teacher has explained but fill in gaps that are blocking tutee comprehension is probably more difficult than providing a full sequence of instruction that the tutor has been preparing to present during the learning part of the Learning-Tutoring Cycle. Not only is a full sequence of instruction more easily delivered than is remediation, but also there is the added benefit that in providing the full sequence, the tutors are much more likely to learn the work thoroughly themselves. Evaluations of the Learning-Tutoring Cycle programs will perhaps indicate if this kind of tutoring does in fact produce stronger gains for tutors and tutees than the ad hoc remediation approach of most existing programs.

Field Studies with Elementary School Tutors

The review of representative secondary school programs ended with a discussion of differences between ad hoc, remedial tutoring of selected tutees, and the approach taken in the Learning-Tutoring Cycle proposal in which tutors prepare and teach solid sequences in the curriculum to entire classes of elementary students. In the first three elementary school projects to be described below we shall see results from two class-to-class tutoring models (Newmark & Melaragno, 1969; and Ebersole, 1971) and from one remedial tutoring project (Niedemeyer & Ellis, 1971).

Two Intact Class Tutoring Projects

Newmark and Melaragno (1969). A Ford foundation grant enabled an ambitious attempt to be mounted to restructure entire schools into "tutorial communities." The "Tutorial Community Project" aimed to involve all school participants--parents, teachers, administrators and students--in shared planning and decision-making for governance of the school. Extensive use of tutoring was one part of this total program, and was, incidentally, the component which survived longest after withdrawal of external funding. Upper grade classrooms were paired with lower grade classrooms for regular tutoring. The upper grade teachers prepared the tutors and regularly held feedback sessions in which the group as a whole dealt with problems encountered in tutoring. As in the work of the Lippitts, major emphasis was placed upon the establishment of caring, understanding, helping relationships between tutors and tutees.

Klaus provides the following account of cognitive outcomes in the area of reading skills: (page 63)

> In addition to some specific studies of various aspects of the tutorial program, records have been maintained on the reading skills of all students since the beginning of the Tutorial Community Project. Overall, there has not been much improvement in reading at the upper-primary grades that can be attributed to the program. At the lower grades, however, both the immediate and the cumulative effects of tutorial assistance have been pronounced. Below are the year-to-year median percentile scores for grades 1, 2, and 3. The program began with first graders in the 1969-70 school year, expanded to the second grade the following year, and was made school wide in the 1971-72 school year.

	1969-70	1970-71	1971-72	1972-73	1973-74
1st grade	23	34	43	60	60
2nd grade	5	19	28	31	40
3rd grade	5	5	20	44	32

The data presented suggest that the first grade results steadily improved, and we understand that population changes did not appear to account for this improvement. If the changing scores can be attributed to the presence of the Tutorial Community Project--and the 5 percentile median scores of classes not involved in tutoring lend support to this attribution--then the major impact appears to have been at the first grade level. Median scores of cohorts typically declined the following year, then held their own. Scores did not, however, drop to the typical pre-project level of the fifth percentile.

Ebersole (1971). Similar improvements in reading skills have been reported following classroom-to-classroom, upper-grade-to-lower-grade tutoring in another elementary school: Soto Street Elementary School, located in a predominantly Mexican-American area of Los Angeles.

In the Soto Street program, advantage was taken of the fact that the lower grade children spend about an hour less time in school than upper grade children. By having one half of the lower grade children arrive early, tutoring could occur between half an upper grade class and the early-arriving lower grade class. These tutoring sessions were closely monitored by the lower grade teacher. Meanwhile the upper grade teacher could work intensively with the other half of the class. At the end of the day the other half of the lower grade students stayed for tutoring by the upper graders who had not yet tutored. This arrangement involved all students in supervised tutoring. Tables 4 and 5, reproduced from Ebersole (undated) show how the distribution of reading scores in 1st and 2nd grades changed as more and more tutoring was gradually undertaken.

Table 4

FIRST GRADE STANFORD READING TEST RESULTS

Nat. Stanine	PRIOR TO TUTORING May Jan May Jan '66 '67 '67 '68	EXPER. TUTORING May 1968	BROADER TUTORING Jan & May 1969	FULL TUTORING May 1970
9				6.5
8			4.2	2.2
7			2.1	2.2 (41%)
6		4.2	6.2 (35%)	
5	1.1 (2%)	4.2 (25%)	4.2	6.5
4	1.1	16.7	18.8	24.0
3	3.5	12.5	27.1	11.0
2	39.8 (94%)	41.7	31.2	41.1
1	54.5	20.8	6.2	6.5

Column 1. In four testing periods* prior to tutoring, the state mandated Stanford Reading Test ranked 94% of the first grade scores in national stanines 1 and 2 while only 2% ranked in the average range or better--stanines 4 and above.

Column 2. The following semester, with experimental tutoring for part of one first grade class using tutors who were in remedial reading classes themselves, 25% of the first grade scores of the school ranked in stanines 4 and above.

Column 3. With broader tutoring and improved tutoring procedures, the combined scores of the next two semesters ranked 35% in stanines 4 and above.

Column 4. The following year, with full tutoring and improved procedures, this percentage increased to 41%, with 6% ranking in stanine 9.

Table 5

SECOND GRADE STANFORD READING TEST RESULTS

Nat. Sta- nine	PRIOR TO TUTORING May Jan May Jan May Jan '66 '67 '67 '68 '68 '69		SOME TUTORING May 1969		FULL TUTORING May 1970	
9					2.8	
8			4.2		2.8	
7	1.7		4.2		5.6	(70%)
6	2.6	(15%)		(54%)	5.6	
5	5.2		25.0		22.2	
4	6.0		20.6		30.6	
3	7.0		8.4			
2	34.4	(77%)	29.2		16.7	
1	43.0		8.4		13.9	

Column 1. In six testing periods* prior to any tutoring, 77% of the second grade scores ranked in national stanines 1 and 2, and only 15% ranked in the average range or better--in stanines 4 and above.

Column 2. The following semester, as a result of some experimental tutoring as first graders and some tutoring as second graders, 54% of the scores ranked in stanines 4 and above.

Column 3. The succeeding year with full tutoring and improved tutoring procedures, the percentage ranking in stanines 4 and above increased to 70%.

Presumably gains were most pronounced for first and second grade classes since these are the results presented, a finding that would echo the Newmark Melaragno results.

We strongly suspect that reading tests are more sensitive to gains due to instruction in the early grades than in later grades. Once a child has mastered decoding, the reading tests start measuring fluency, comprehension, general knowledge--skills not necessarily easily affected by school instruction.

Both the Newmark and Melaragno and the Ebersole programs suggest that classroom-to-classroom tutoring, aiming at consistent daily instruction rather than piecemeal remediation produced excellent results. We examine now a report of tutoring used for remediation.

More on Ad Hoc Remediation

Niedermeyer and Ellis (1971). The developers of individualized learning programs often find themselves becoming involved in cross-age tutoring, since it provides one of the few feasible means of individualizing instruction without demanding that the younger learner be his or her own teacher. The project reported by Niedermeyer and Ellis grew out of product development activities of the Southwest Regional Laboratory for Educational Research and Development (SWRL). Their interest was in the use of tutors to improve tutee performance in reading. The project they described was developed and tried out, in the course of several school years, in schools attended by children from middle-income families in suburban Los Angeles.

Tutors were fifth and sixth grade pupils. Their method of selection was not described, but they left their regular classes to tutor. Tutees were kindergartners who "needed remediation." A kindergartner needed remediation for a reading unit if he or she achieved less than 18 out of 20 on a test for that unit of the reading program.

The authors refer to the tutors' work with the kindergartners as "monitoring." The training of tutors, which was developed after summer work with tutoring, emphasized seven instructional techniques relating to feedback and reinforcement. The content the tutor was to work with was a practice exercise prescribed by the kindergarten teacher. Since it was reported that kindergarten teachers assigned the tutors "by pinning a tutor's name on the Tutor Assignment Sheet," it seems likely that tutors were assigned to work with various tutees on an ad hoc basis, as needed. Their responsibility was to the teacher's assignment rather than an on-going one for a particular child. Three times a week tutors worked with tutees for 20-minute sessions. They collected their assignment, completed the Practice Exercise, and might then be assigned to "monitor" the reading of a storybook or might take the tutee out to play.

Data were presented to answer two questions: Did the tutoring benefit tutees, and did the training and/or experience of tutoring alter the behavior of tutors?

To test whether tutoring benefitted tutees, the authors randomly selected four of eight schools to use the tutorial program. In the four remaining schools, teachers managed the remediation activities without the assistance of fifth and sixth grade tutors. Four remediation-needing children were randomly selected from each of seven classes in each set of schools, and the criterion test was administered by personnel of SWRL. The authors claimed "a difference of 14 percent in favor of the classes with tutors" on the basis of the gains shown in Table 6, reproduced from Niedermeyer and Ellis.

Table 6

Niedermeyer and Ellis Results
Regarding Effectiveness of Tutors:
Percent of Pupils Scoring Above 90 Percent
on Unit Tests and on Retests Following Remediation

Remediation by Teacher Only				Remediation by Teacher Plus Tutor			
Teacher	Test	Retest	Percent Gain	Teacher	Test	Retest	Percent Gain
1	81	93	+12	8	69	86	+17
2	57	86	+29	9	57	67	+10
3	85	91	+ 6	10	68	92	+24
4	51	16	-35	11	55	62	+ 7
5	55	80	+25	12	70	93	+23
6	78	81	+ 3	13	53	76	+23
7	91	91	0	14	56	68	+12
Mean	66	69	+ 3	Mean	62	79	+17

The authors' claim appears highly dubious in view of the unexplained outlier which was left in the control group; one group showed 35% fewer pupils scoring above 90% on the posttest than on the pretest. Removal of this strange result and a re-analysis of the data reveals results favoring the control, non-tutored classes.

The Niedermeyer and Ellis study is important in that it examined the question as to whether elementary school classes benefitted from having upper grade tutors available for remedial work. Did classes do as well without the tutors? It seems that they did.

Effects of tutoring on the tutors were not assessed, but it is worth noting here that students generally prefer consistent responsibility for their own tutee or tutees rather than daily assignments to whomever needs help. For example, reporting from a school that had both teacher aide and tutoring programs, Klaus noted, " . . . students generally prefer the program with more training, more weekly support, and more opportunity to work with a single child over a long period." Fitz-Gibbon noted that even tutors who complained of difficult-to-teach or difficult-to-manage tutees did not wish to change the tutees.

To summarize this discussion: the possibility should be considered that there are the following advantages to assigning full-sequence instruction rather than remedial work:

(i) it is easier for tutors to be effective in full-sequence instruction. Clearly children selected because they are already having trouble with the work are more difficult to teach than other children so that tutors have less chance of seeing success when assigned only to remedial work. The fact that tutees are already confused makes the diagnosis and prescription more difficult. (This is not to say tutors should not teach difficult children. If these same children had been presented the work initially in a tutoring situation, they might have paid closer attention and have avoided reaching a point of confusion.)

(ii) Tutors may learn the work more thoroughly when going through the process of full-sequence instruction rather than handling only specific remediation. Full-sequence instruction demands explanations, illustrations, logical development and comprehensive coverage. This processing, as was suggested earlier in this review, might be highly important in tutor gains.

(iii) Full-sequence instruction avoids an association of tutoring with failure. If everyone in the class is tutored, there is no encouragement to the tutee to resist tutoring or dislike it because it is a stigma, indicating one has failed in some way.

(iv) Tutors may prefer full-sequence instruction. Being assigned different tutees every few days and told to teach bits and pieces of the curriculum may lead to disgruntled tutors. (Much will depend on the manner of the tutee's teacher, of course.)

The studies just cited do not provide a basis for drawing final conclusions regarding the advisability of using tutors for remediation as opposed to full-sequence instruction. Suffice it to say that the issue is alive and should be the subject of further investigation: remedial tutoring seems to be the first instinct of practitioners but it may not be the best.

The next study to be described illustrates full sequence instruction with fifth and sixth graders teaching second and third graders.

A Learning-by-Tutoring, Full Sequence Intstruction Project

Mohan (1972). Mohan reports a study of the effects of tutoring on unmotivated students--both tutors and tutees. Students who, on the basis of assessment by themselves and their teachers, were in the lowest 10% for motivation were randomly divided to form experimental and control groups. The control group students remained in regular classes whereas experimental group students participated in two hour-long tutoring sessions per week. Tutors were fifth and sixth graders and tutees were second and third graders. There were 16 pairs tutoring and 31 control subjects. The tutoring covered set theory for second and third graders.

Tutors were assigned to teach specified content areas each day, for which they were instructed to prepare lessons. Four test items keyed to the content area were provided to assist the tutor in diagnosis. In addition to teaching, tutors were encouraged to talk with their tutees about study habits, test taking, information searches, manners, success, and appearance. The project was run for three months.

In Table 7 posttest mean scores for tutors and the relevant control groups are displayed along with univariate F-tests. Similar data for tutees are presented in Table 8.

Table 7

Means and F-test Between the Post-test Scores
of Control and Experimental Tutors

Variable	Control Group Means	Experimental Group Means	F	P
Mathematics	16.76	31.00	163.6362	.0001*
Motivation (T)[1]	35.88	43.69	13.8117	.0011*
Motivation (S)[2]	48.17	51.81	2.3296	.1395
Attitude	44.76	54.19	4.2373	.0502*
Self-Concept	49.71	54.75	.0023	.9618

*Significant at .05 level.

1 stands for motivation as assessed by the teacher

2 stands for motivation as assessed by the student

Table 8

Means and F-test Between the Post-test Scores of Control and Experimental Tutees

Variable	Control Group Means	Experimental Group Means	F	P
Mathematics	7.79	24.44	85.7570	.0001*
Motivation (T)[1]	38.86	41.38	1.4619	.2308
Motivation (S)[2]	43.79	47.69	5.2312	.0322*
Attitude	20.07	18.50	.1067	.7471
Self-Concept	23.85	24.06	.0377	.8490

*Significant at .05 level.
1 stands for motivation as assessed by the teacher
2 stands for motivation as assessed by the student

(Tables 7 and 8 reproduced from Mohan, 1972)

The mathematics test was locally developed to assess the objectives in set theory. Clearly, tutors and tutees knew the work substantially better than the randomly constituted control groups. It is not clear from Mohan's account whether tutoring replaced some normal classwork or was an add-on activity, i.e., the influence of time-allocated was not necessarily controlled for in this study. However, tutors clearly learned from teaching, and tutees learned from being tutored.

Tutors Gain When the Task Matches Their Needs

The last study to be examined in this section illustrates one of the principles most requiring emphasis when the effects of tutoring projects are under consideration: Immediate, measurable cognitive gains by tutors can only be expected if they are tutoring in a topic they need to learn or practice. Tutors learn what they teach. If what they are teaching is work they already know thoroughly, then learning gains will not ensue. (The reader will recall that it is the selection of the tutoring task to match the needs of the tutors which distinguishes a Learning-by-Tutoring project from other kinds of tutoring projects.)

A number of projects have selected the tutoring tasks to meet the needs of the tutees while nevertheless expecting general learning gains from tutors. There are three possible reasons why tutors might show learning gains in such a project. First, tutors' needs may quite fortuitously coincide with tutees' needs thus resulting in learning gains because tutors are then teaching work they need to learn. A second possibility is that by teaching, and by observing the learning process in tutees, tutors "learn how to learn." This possibility was strongly suggested by an influential book (Gartner, Kohler and Riessman, 1971) but evidence for such a process has not been produced. A third way in which tutors might show general learning gains is through "non-cognitive" effects such as improved self-esteem or improved academic motivation which are assumed to lead to general academic improvements. Like "learning how to learn," this improved-achievement-deriving-from-non-cognitive-effects hypothesis is appealing but unsubstantiated by evidence. It seems reasonable to expect that even if the latter two processes (learning-how-to-learn and non-cognitive effects on learning) did occur, they would require projects of either long duration or high intensity to yield measurable effects. More research is needed before the existence of either of these processes can be accepted, but the evidence for the more modest hypothesis that tutors learn what they teach is now available from many studies.

The study which provided an excellent illustration of the principle that tutors show cognitive benefits primarily when the tutoring content matches their needs was conducted by Olds (1976), a student of Urie Bronfenbrenner. Olds was able to contrast being a tutor with receiving tutoring from an adult. Sixth graders who tutored second graders for three months made greater vocabulary gains than sixth graders given their own adult tutors during the same time. Olds attributed this to the "initial enthusiasm and preference for tutoring as opposed to being tutored." At the end of a further three months, disordinal interaction was located on the assessment of gains in reading comprehension: better readers made greater gains when they received tutoring from adults, whereas poor readers made greater gains when they acted as tutors to younger children. Olds commented, "Poor readers needed basic skill review, and therefore made gains as a result of teaching basic skills to second graders. Better readers, however, needed work in more advanced areas of reading: teaching elementary skills added nothing to their repertoire of reading abilities."

As emphasized above, cognitive gains for tutors should not be expected unless they are teaching work which they themselves need to learn or practice.

More on the Issue of Standardized Tests

With the exception of the Cloward (1967) study, the tutoring projects examined in this literature review have been those reported in 1969 or later. Rosenshine and Furst (1969) undertook a survey of objective data available from Educational Resources Information Center (ERIC) files through Spring, 1969. A summary of this survey was presented at a conference of the American Educational Research Association and has been widely quoted. Unfortunately, they took an approach to the evaluation of interventions which involved a misplaced faith in grade equivalent scores. For example:

> In most studies, the results were frequently not presented as grade equivalent scores, so that it is difficult to assess the effectiveness of tutoring in any educationally significant way. In other words, the experimentals "beat" the controls, but as far as we know, both may have been far below grade level at the conclusion of the projects (p. 3).

Their concern was a laudable one: to ensure educational and not just statistical significance of results; but their solution, grade equivalent scores, is not satisfactory. As noted in CSE Report on Tutoring #1 (p. 58 ff), grade equivalent scores are frequently interpolations and extrapolations of dubious validity and should not be used as criteria (see also, for example, Mehrens and Lehman, 1969, p. 55; Lyman, 1971, p. 115).

The uninterpretability of grade equivalent scores was well illustrated, in fact, in the Rosenshine-Furst review. They themselves noted that, from the gains reported by Cloward for his control group, it appeared that simply by giving a retest one could raise the achievement level of high school students who are reading six months below grade level, up to grade level and beyond.

Immediately after noting that suspicious result, they report a pre-post difference which was not statistically significant but which showed a year's gain in grade equivalent scores (Grannick, 1968). Placing more faith in grade equivalent scores than in a statistical test applied in the easiest situation for showing a significant difference (a single pre-post comparison), they tended to view the result positively:

> Grannick's Philadelphia group had a mean of .4 years below grade level on the pre-test. They gained one year during a seven week period. These gains, however, were not statistically significant. One might, however, argue for educational significance of a finding which shows grade level attainment in a short period (p. 6).

Rather than urge the translation of raw scores into scores with misleading labels (grade equivalents) better advice to researchers and evaluators would be to implement good research designs, provide competitive control groups that receive the same allocated instructional time and use sensitive measures of student achievement, such as tests referenced to instruction. This advice should inform the work of evaluators of tutoring projects.

It is of interest to note that the list of "unsuccessful" projects presented by Rosenshine and Furst consisted almost entirely of reading projects. (The single exception was one of their own studies which was reading and math. They undoubtedly used standardized tests as the measures.) The difficulty of measuring reading achievement is considerable. Before reading is established, results will be strongly affected by method of teaching employed (e.g., systematic or intrinsic phonics); and once reading is established, reading tests become thinking and reasoning tests that may be insensitive to instructional effects. Projects in which math achievement is examined might yield more consistent results because math achievement may be more influenced by school instruction. (See, however, p. 79 below.)

In discussing the "unsuccessful" projects, Rosenshine and Furst located a possible reason for lack of significant findings as residing in the task charcharacteristics:

> ...the unsuccessful programs were more diffuse in their objectives, attempting to provide a variety of enrichment services. In the successful programs, more time was spent on cultural activities, and less time on academic activities" (p. 4).

(Ellson, Harris and Barber, 1968, advanced a similar explanation for the superiority of programmed tutoring over traditional adult tutoring.)

One must wonder if the programs were "unsuccessful" in all aspects, not just in cognitive gains. Highly structured programs focus time on cognitive tasks. Are attitudinal benefits thus neglected? Do attitudinal benefits need specific plans and activities, or do attitudinal benefits simply occur in conjunction with successful cognitive benefits? These are just a few of the questions which must still be answered regarding cross-age tutoring programs.

Summary of Cognitive Outcomes

Several controlled experiments have examined the cognitive impact of tutoring with positive results. Fitz-Gibbon (1975) showed that low achieving, inner-city, ninth grade tutors learned more during three weeks spent tutoring fourth graders than an equivalent control group learned by spending the same three weeks working on the same objectives in their regular classes. The positive effects of the tutoring experience were still significant three months after the project. Allen and Feldman (1973), in a summer school project with paid participants, found fifth graders eventually achieved higher scores when assigned tutoring than when assigned the same time to study alone.

Is it the expectation of tutoring or the process of tutoring that produces the cognitive benefits? Present limited evidence, it was argued, suggests that the process of tutoring is the more influential variable, which implies that the *act of tutoring makes a special contribution to learning, one that might not be effectively replaced by other means of motivation.*

The effect of tutoring on tutees was examined by Bausell, Moody and Walzl (1972) contrasting tutoring with equivalent time spent as a member of a class. Results favored being tutored over receiving classroom instruction from the same teacher. An experiment by Marks, Doctorow and Wittrock (1975) was noted because it lends credence to the idea that the closer language-match between tutor and tutee might be one of the advantages that peer-tutoring has over adult tutoring.

Results obtained by Hamblin & Hamblin (1972) which indicated that teenage tutors were not as effective with kindergartners as were kindergarten tutors were rejected as unsubstantiated. An experiment (Snapp, Oakland and Williams, 1972) in which upper grade students tutored lower grade students before school reported significant gains for tutees on a test keyed to instruction but no effects on a standardized test. *Throughout the review of the literature, the congruency between outcome measures and the specific tutoring activities is crucial in understanding the results. Students learn what they do, but diffusion and generalization of these gains either do not occur or are difficult to measure.*

Following examination of controlled experiments, field studies--programs implemented in natural settings for extended periods of time--were examined. Cloward's (1967) study of a Youth-Tutoring-Youth program showed significant gains for tutors over the course of a year's tutoring. Tutees showed significant gains only when tutored for two sessions a week. In effect, one hour a week of academic tutoring was not sufficient to produce significant gains, but two hours a week made a significant difference. Projects modeled after the Ontario-Montclair cross-age helping programs were described. These are flexible programs in which secondary tutors, who take tutoring as an elective, work with children pulled from various classrooms at the elementary school. Emphasis is on the helping relationship and evaluation of cognitive outcomes would be difficult in view of the variety of activities undertaken by tutors.

A more structured classroom-to-classroom cross-age tutoring program was described (Hoffmeister, 1973). Ninth graders taught seventh graders in a remedial reading class in a Title I school with results for both groups that would be classified as successful by the usual Title I criterion of month-for-month growth. The Hoffmeister program was approximately a Learning-by-Tutoring project implemented within a junior high school. The extreme reading-retardation of the tutees was noted as a possible reason why tutors did not make greater gains.

A project reported by Dreyer (1973) emphasized the need for an effective Tutor Support System.

The review of those field studies in which tutors were secondary school students ended with the question as to whether assigning tutors to conduct remedial work was less desirable than assigning them to provide "full-sequence" instruction. Looking at projects in elementary schools, strong results were found for two projects involving regular, full-sequence, classroom-to-total-classroom instruction [Newmark and Melaragno (1969) and Ebersole (1971)] but unconvincing results were noted for the benefits to tutees of selective remedial work by upper grade tutors [Niedermeyer and Ellis (1971)].

Arguments advanced in favor of full-sequence instruction were that it makes teaching success for tutors more likely, promotes better tutor learning, avoids the stigma associated with individual remediation, and that tutors are likely to prefer responsibility for full sequence instruction with a single child to piecemeal remediation activities with a succession of children.

Mohan (1972) reported a field experiment in which upper grade elementary students tutored lower grade elementary students in set theory. Gains were strong and significant for both tutors and tutees in comparison with randomly constituted control groups remaining in regular classes. A test closely keyed to instruction, it should be noted, was the achievement measure employed by Mohan. Results of a recent study by Olds (1976) supported the principle that to derive cognitive benefits, tutors must teach work they need to learn or practice. A review of pre-1970 projects by Rosenshine and Furst was discussed with particular emphasis on the misleading nature of grade equivalent scores. The relevance of assessment to instruction was seen as an important feature determining the results of studies.

CHAPTER

CHAPTER III. COGNITIVE OUTCOMES IN PROGRAMMED TUTORING

The literature on cross-age tutoring is generally highly positive. Perhaps the most negative note that is sounded is by some proponents of programmed tutoring who have contended that non-programmed tutoring is ineffective--meaning that tutees do not learn. It must be noted immediately that programmed materials are, almost by definition, effective. Many years of development goes into ensuring that the materials, when correctly administered, achieve their objectives. We shall take as given, then, the proposition that materials that are programmed to achieve objectives do achieve them. The question is whether non-programmed methods can produce sufficient gains for most students to make ordinary tutoring worthwhile or whether tutors should all use programmed materials.

Effects on Tutors of Administering Programmed Tutoring

Before describing some studies that have examined the effects of programmed tutoring on tutees, the question of tutor benefits must be raised. Klaus reports that student tutors rapidly became bored using highly structured materials (Klaus, 1975, p. 53). It is important to note that most programmed materials have been developed and implemented using paid, adult tutors, a situation in which tutor motivation has not been a primary consideration.

A couple of experiments speak to the question of cognitive benefits to tutors who administer programmed materials.

Myers, Travers, and Sanford (1965). These researchers examined the performance of fourth, fifth, and sixth grade pupils on a learning task when the pupils were randomly assigned to one of four learning conditions. The four learning conditions, three of which involved students working in pairs on the learning task, were: Condition I: tutor role; Condition II: tutee role; Condition III: tutor-tutee switching; and Condition IV: self-instruction.

The learning task was rote memory for the English translation of German nouns. The German noun and two choices for its meaning were printed on cards. In the tutor role the student showed the card to the other pupil, and on the back of these cards, as they were held up, the three words were displayed with the correct English choice indicated. The tutor role students could therefore see the correct answer as soon as they posed the item to their partners. The students in the tutee roles were to guess the correct answer, after which the tutor stated the correct answer without comment. For the tutor-tutee switching, students reversed roles in mid task. In the self-instruction role, students studied alone from cards. In all conditions, students were allowed whatever time they needed to proceed through a set of cards. Amount of time was generally 15 to 20 minutes. The procedure was repeated with new word lists, i.e., new tasks, on three consecutive days. Results for the posttest data indicated significant effects for conditions and tasks and a significant interaction between tasks and conditions. Figure 9 is reproduced from the article (p. 70).

The steep decline across the three tasks (representing three consecutive occasions) was taken as indicative of boredom with the materials. The students in Condition I, the tutor role, achieved the lowest scores both at immediate posttest and for retention tests given from three to nine weeks later. We should note that holding up a card and reading what is written on the card is not what is generally thought of as "teaching" or "tutoring," but it may constitute a role similar to that of the tutor who is administering drill or programmed materials.

The Myers, et al. experiment suggests that little benefit will accrue to tutors who essentially take the place of a teaching machine. The difference between tutors and tutees in this experiment was essentially that tutees had to confront the material and make a choice for his or her guess. Apparently this amount of active processing gave the tutee an advantage over the tutor who was passively exposed to the same materials.

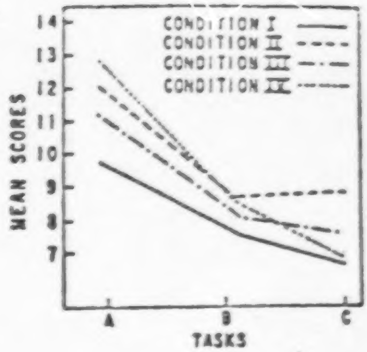

Figure 9. Myers, Travers, and Sanford results: Graphical representation of the interaction of tasks and learning conditions.

Bull and Wittrock (1975). Further support for the contention that, in order for tutors to benefit, they must be actively engaged in teaching--that is, in processing, presenting, and organizing the material for the tutee--comes from the work of Wittrock and associates. They are interested in learning as a "generative process," one in which the learner must be actively involved.

Ninety fifth graders were randomly divided into three groups, each of which was instructed to learn word definitions in one of the following ways: (1) generate (draw) an image of the word and its definition, (2) trace a picture (image given) representing the word and its definition, or (3) learn the verbal definition by copying it. The group that remembered the definitions best one week later was the group that generated their own images. Bull and Wittrock commented that, considering this and other studies, what seems to be important is that some type of generation activity is called for.

It seems likely that tutors who know what they have to teach (like those pupils in the experiment who saw the word definitions that they had to learn) will retain the work better if they have to generate their own way of teaching it (as the subjects in Bull and Wittrock's experiment had to generate their own images) than if they simply follow prescriptions for teaching (as did subjects in the experiment who simply followed instructions for tracing provided images).

Ronshausen, developer of programmed tutoring materials in mathematics, emphasized the minimal demands on tutors:

> The need for professional judgments by the tutor is reduced to a minimum--the tutor judges only the correctness of the child's responses. While the operational programs describe correct and incorrect responses and tell the tutor how to proceed in either case, the content programs give specific correct responses. Thus, persons with limited knowledge of mathematics can perform successfully as tutors (Ronshausen, no date).

Similarly in the evaluation of a Title I project in Highland Park, Michigan, the role of tutor is seen as one of replacing a machine:

> Tutors in this program easily performed functions which if performed by teaching machines would have entailed great expense. They kept records of correct and incorrect responses, delivered immediate reinforcement for correct responses and, most important, provided the correct response whenever an incorrect response occurred (Highland Park School District, 1976).

Both the Myers et al. and the Bull and Wittrock experiments suggest that tutors will certainly not derive maximal benefits from administering programmed materials. (This is not a criticism of programmed materials, because they are designed for tutees not for tutors, but a warning for planners of tutoring projects.)

We turn now to an examination of two programmed tutoring projects. There have been two very thorough product development activities related to tutoring, one in the area of elementary school reading (Ellson, et al., 1968) and one in primary grade mathematics (Ronshausen, 1974). The focus in both these projects, which ran for several years, was benefit to tutees. Tutors were usually paid paraprofessional adults.

Programmed vs. "Directed" Tutoring for Beginning Reading

Ellson, Harris and Barber (1968). Ellson et al. reported a 28-week field test in 20 elementary schools in Indiana, 10 of these schools being in the inner city. The field test implemented a 2 x 2 x 2 design with the following factors:

- Tutoring versus no tutoring
- "Programmed" versus "directed" tutoring
- One versus two sessions of tutoring daily

"Programmed" tutoring referred to a set of "operational programs" and "content programs" which tightly controlled the tutors' procedures. These programs had been developed over several years of experimental work. "Directed" tutoring consisted of procedures recommended by a reading specialist and reading consultants; it was described as an adaptation of classroom methods to an individual situation. As such, Directed tutoring employed a wide variety of support materials such as readers, pictures, games, crayons, etc.

Selection and characteristics of tutors and tutees. Tutors were recruited by notices in school and Parent-Teacher Association bulletins. All were white and female and had had at least a high school education. They were paid on an hourly basis.

Since an important question to be answered by this field test was the relative effectiveness of Programmed versus Directed tutoring, it was a serious error that tutors were not randomly assigned to these methods of tutoring:

> The selection of tutors was subjective, based largely on the judgment that the candidate was capable of working well with children and school personnel and competent to carry out the required tutoring and recording procedures. Assignment of tutors to programmed or directed tutoring was presumably influenced by judgments of compatibility between personality traits and tutoring tasks, but the factors which entered into these judgments or their actual relevance is not known. Nine of the programmed tutors were carry overs from the previous experiments, i.e., they had one semester experience in programmed tutoring (p. 310).

Thus, not only was the assignment of tutors non-random, but the Programmed tutoring treatment was confounded with experience. A newly developed, first trial, Directed tutoring program, employing persons with no prior tutoring experience was compared with tutoring which had had several years of revisions and which was implemented in part by tutors with some prior experience. (The total number of tutors was not reported so the proportion of experienced tutors cannot be ascertained.)

Tutees were a representative sample of first graders in Indianapolis schools located in lower, and lower middle-class districts. Sixty children were assigned to each experimental group by "arbitrary procedures designed to equalize groups and give a representative sample of the first-grade population in the schools assigned to the study" (p. 323).

Training of tutors. Tutors received 12 hours of training (four sessions of three hours each) prior to commencing their tutoring activities. A further nine hours training was given during the program (in three three-hour sessions) and there was home study and close on-the-job supervision during the first few weeks.

Sessions. Tutoring was done outside the regular classroom in 15-minute sessions. If tutees received two sessions a day, these sessions were separated by at least an hour. It should be noted that tutoring represented additional time of up to 30 minutes per day devoted to tutoring. Description of the activities of each session gave a very favorable impression of both the Programmed and Directed tutoring.

Regarding Programmed tutoring the authors wrote:

> For one who briefly encounters a child being tutored, the fact that there is a program probably would not be visible at all: the materials and

> the behavior of the child and of the tutor sitting beside him are well within the scope of traditional teaching practice. With experienced tutors, the process appears quite spontaneous, but after a little time it becomes apparent to an observer that the tutor's behavior is highly systematic. It would almost certainly be noticed that the tutor makes a more detailed record of each child's behavior, specifically his successes and failures, than is ordinarily made, although it might not be immediately evident that almost every act and word of the tutor is very tightly controlled by the recorded information.
>
> After much extended observation, the nature of the system might be discerned, but its description would likely vary. An experimental psychologist might see it as a complex and flexibly modified form of the paired associates anticipation procedure, supplemented by verbal approval as a form of reinforcement. A programmed instruction specialist would notice the frequent and immediate reinforcement of correct responses, the loops or branches in the program, the use of prompts, and the fact that the material to be learned was divided and taught in small units. An experienced teacher would see many examples of her art systematically employed and also that certain common practices of teachers were systematically avoided. For example, in its present form programmed tutoring of reading does not include and is not preceded by a readiness program and no special devices beyond the learning situation itself are used to motivate and stimulate the children's interest. These appear to be far less necessary in programmed tutoring than in group teaching. It should be remembered, however, that activities designed to develop readiness and interest-oriented motivational techniques are likely to be included in the work of the classroom that is concurrent with the tutoring.
>
> An educator would perhaps recognize the programs as simple forms of the discovery method. The first step in each program presents a problem or task in its most difficult form with a minimum of context. If the child cannot solve this form of the problem, it is progressively simplified by changing its form or by providing hints or additional information until the child discovers the solution for himself. (p. 314)

Regarding Directed tutoring the authors wrote:

> . . . those who observed it in operation during the course of the experiment had every expectation that the participating children would benefit. The sequence of directed tutoring activities was planned in detail by an experienced reading specialist with the guidance and support of several experts.
>
> The training of the tutors appeared to be adequate for the limited task involved and was well supported by supervision which provided additional on-the-job training when necessary.
>
> Both tutors and those who observed them had no doubts that a real contribution was being made to the children's reading skills. (p. 343)

Evaluative data. Table 9 was drawn up from data presented in several tables in the article.

Table 9

Ellson, Harris and Barber Results:
F-Ratios and Significance of Main Effects

	Analysis Employing Total Group		Analysis of Tutored Groups Only					
	Tutoring vs Non-tutoring		Programmed vs Directed		1 session vs 2 sessions		Inter-actions	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Midtests								
PPVT	4.99	<.05	3.96	<.05	.01	n.s.	Not reported	
Ginn Total	4.90	<.05	4.27	<.05	1.66	n.s.		
Posttests								
PPVT	3.33	n.s.	.03	n.s.	.89	n.s.	.10	n.s.
Ginn Total	3.34	n.s.	24.46	<.01	.93	n.s.	4.39	n.s.
Stanford Reading	.51	n.s.	4.26	<.05	.69	n.s.	.98	n.s.

Significant results at midtests favoring "tutoring" over "no tutoring" and Programmed tutoring over Directed tutoring were sustained at posttest time only by Programmed tutoring which excelled over Directed tutoring.

When each separate tutoring condition was contrasted with control group scores, the results were as shown in Table 10.

Table 10

Ellson, Harris and Barber Analyses of Covariance:
F-Ratios and Significances of Experiment vs. Control Effects
on Posttests based on Analyses of Covariance,
Each Tutoring Condition is Analyzed Separately[a]

Posttest Measure	Condition							
	Programmed tutoring				Directed tutoring			
	1-session		2-sessions		1-session		2-sessions	
	F	p	F	p	F	p	F	p
PPVT	1.83	n.s.	.03	n.s.	6.69	<.05	.01	n.s.
Ginn Total (A + B)	.01	n.s.	12.03	<.01	1.67	n.s.	.70	n.s.
Stanford Reading Total	.00	n.s.	1.09	n.s.	2.56	n.s.	1.03	n.s.

[a] The degrees of freedom were 1 and 83. The F-ratio needed at the .05 level of confidence was 3.95 and 6.94 for the .01 level.

The authors note that the lack of significant differences in the Stanford Reading tests "may reflect one of the difficulties of obtaining a satisfactory evaluation of reading achievement at the beginning level. The Ginn tests are aimed directly at the content specifically taught through the Ginn Basal Reader Series. The Stanford test samples a broader content which only partially overlaps that of the Ginn Series" (p. 334).

The results indicated that significant differences were the result of high scores of tutees who had received two sessions of programmed tutoring per day. Only this treatment appeared to be effective and this only on the Ginn test. On this test there was a very substantial effect, presumably due to tutoring. The authors noted that "significant differences were limited almost entirely to measures highly weighted with vocabulary content" (p. 333). The significant effect for Directed tutoring in the Peabody Picture Vocabulary Test (PPVT) was not discussed.

As already stated, the development of programmed procedures, if based on empirical validations throughout successive iterations and revisions of the procedures, must result in an efficient, effective set of learning procedures. Consequently, positive results from the Programmed tutoring were strongly to be expected. The nonsignificant effects on the Ginn total and on the Stanford Reading total for one session per day were therefore surprising, especially since the session of Programmed tutoring represented an addition of 15 minutes per day to regular instruction for 28 weeks.

The lack of effects for either one or two daily sessions of Directed reading are also surprising when one has read of the apparently well-financed, well-conducted tutoring procedures.

The problem could be the measuring instruments. Giving group-administered standardized tests to first graders is not necessarily a valid measurement procedure. At this age, the willingness and ability to follow instructions and concentrate may have more influence on the outcomes than the reading skills purportedly being measured. Many children at this age tend to ask examiners if their answers are correct, share their answers with other children, or abandon the task completely, problems leading to a questionable validity for the measures.

Furthermore, the complexities of the problem of measuring reading achievement cannot be underestimated. If the tutoring emphasized a systematic phonics approach to reading skills (and it does appear to have done so), then it could be argued that one should not expect the benefits to be apparent in first grade. To quote a researcher who made an extensive study of reading methods:

> The phonics-trained children, especially those given more systematic-phonics instruction, acquired a set for accuracy in recognizing words and were preoccupied with letters and sounds (as noted in Currier and Duguid, 1916; Winch, 1925; and Gates and Russell, 1938). This kind of learning set pays off in oral, untimed tests of word recognition, but can lead to lower scores on silent, timed reading tests, especially in grade 1, when the average phonics-trained child has not yet acquired sufficient skill to "sound out" words that he cannot recognize immediately. This child concentrates on working out the words and has less time and energy left to devote to getting the thoughts. By the second grade, however, when the average phonics-trained child has mastered a sufficient amount of phonics knowledge and skill, he can devote himself more to meaning. (Chall, 1967, p. 107)

There is also the question of the fairness of the tests to the two kinds of tutoring, Programmed vs. Directed.

The Ginn test on which a significant result was obtained was apparently more closely tied to the instruction which had been received than was the Stanford. The problem of test-fairness arises because if the outcome measure which showed Programmed tutoring superior to Directed tutoring were a test tied to the vocabulary built into the Programmed tutoring, but not into Directed tutoring, this would hardly be a fair test. Ellson developed the Programmed tutoring over several years, with many changes (p. 346), and has written at some length on the Ginn tests. A link between program and test cannot be ruled out. One must wonder if some tutoring activities were highly test-related, such as by the presentation of similar vocabulary or item formats. Even if we set aside the suspicion that the Ginn test might have favored the specific content found in Programmed tutoring, there is still the problem already pointed out that tutors were not randomly assigned, nor equally experienced and the Directed tutoring was receiving its first tryout.

However, the authors offered an explanation for the better results from the Programmed tutoring which appears both highly plausible and interesting. They point out that the Directed tutoring involved many "support" activities, whereas in the Programmed tutoring the time was totally devoted to reading.

> It is possible that the relative effectiveness of the two types of tutoring has a simple explanation. Harris and Serwer (1966a, b) found that of many variables operating in the classroom, the amount of time that the teacher devoted to reading activities was among those most highly correlated with reading achievement, while the correlations for amount of time spent in supportive activities were slightly negative, but not significant. Thus, since Programmed tutoring is entirely devoted to reading activities, the difference in their effectiveness may simply reflect the extent to which the two techniques supplement the classroom with additional time devoted to the teaching of reading. (p. 342)

This seems a reasonable explanation, given the development methods that produced the Programmed tutoring procedures. *The concern which must be expressed is that the results concerning Directed tutoring not be overinterpretted to the point of stating that tutees do not learn unless tutors follow high structured, externally imposed procedures. Such a conclusion was not established by the Ellson, Harris and Barber study.* The lack of significant effect for one session of Programmed tutoring lends support to doubts concerning the sensitivity of standardized reading tests: surely an extra 15 minutes per day of reading instruction with effective materials should produce results.

Finally it should be reiterated that the issue of "Programmed" versus "Directed" tutoring should be resolved solely on the basis of the outcomes for tutees, since benefits for tutors are also frequently a goal of tutoring projects.

Ronshausen (1974). As with the Ellson et al. project, Ronshausen's project involved long-term testing and development of a programmed tutoring system, focusing on benefits to the tutees. Her programmed techniques for the teaching of beginning mathematics involve the same "discovery learning" process described by Ellson et al., a process Ronshausen has termed "brightening." Her tutors, as Ellson's, are trained to avoid negative comments completely and to reward all correct responses with positive reinforcement.*

As with Ellson et al., Ronshausen's tutors were paid adults. Each tutor received about 30 hours of training in the two component programs of the "Programmed Math Tutorial" (PMT). The two programs were the operational program, which defines general tutoring methods (applicable to any subject area like teacher training "methods" courses), and the content program (the actual mathematics to be taught, organized for discovery learning and carefully structured, sequential mastery learning.

---

*The avoidance of negative comments for incorrect answers is incorrectly termed by Ronshausen (as frequently elsewhere), as the avoidance of "negative reinforcement." Technically a negative reinforcement occurs when, contingent on the correct response, an aversive stimulus is discontinued. This discontinuance serves to reinforce (reward) the correct response. What is avoided, in these programmed tutoring prescriptions, is not "negative reinforcement" but "punishment" such as negative comments.

A large field test of PMT, in its third cycle of development, was conducted by Ronshausen in the 1972 to 1973 school year. This field test had two objectives: to test first graders on a revision of first grade PMT materials, and to see if these same materials could be effectively used for kindergartners as well as for the first graders for whom they had been originally developed.

The results are shown in Tables 11 and 12, reproduced from Ronshausen (1974). Note that tutored students (drawn from grades K and 1 in one school system and grade 1 in another school system) achieved significantly higher scores on all tests of mathematical concepts, but not on all tests of computational skills. Ronshausen concluded:

> The use of PMT as a supplement to regular classroom instruction seems to be more effective than regular classroom instruction alone in helping the children learn mathematics concepts. . . . The combination is not particularly more effective in helping first-graders learn computational skills. The lack of significant differences between the mean scores of the tutored and untutored control groups may indicate that classroom teachers devote more time to teaching first-graders computational skills than mathematics concepts, or that their instruction is more effective for computational skills than for mathematics concepts. Another possibility pertains to the design of the PMT--mathematics concepts are emphasized throughout. Computational skills are not neglected, though, and the findings in the kindergarten field study seem to suggest that the lack of a significant difference in computational skills mean scores for first-graders is not due to the design of the PMT. (p. 12)

Table 11

Ronshausen's Summary of Math Concepts Tests:
Mean Scores for Treatments

School System	Grade	Subtest	Treatments Tutored	N	Untutored Control	N	Diff. in Means	t	p
A	1	Metropolitan Achievement Test, Primary I, Subtest 4A	17.2 (17.7)**	140	16.7 (15.9)	85	1.8	2.97*	<.01
A	1	Criterion Items Subtest, Concepts Section	6.1 (6.3)	140	6.0 (5.7)	85	.6	2.55*	<.02
B	1	Metropolitan Achievement Tests, Primary I, Subtest 4A	21.3	136	19.8	78	1.5	2.03	<.05
B	1	Primary School Mathematics Criterion Test, Concepts Section	28.6	136	25.3	78	3.3	4.86	<.001
B	K	Metropolitan Achievement Tests, Primer, Concepts Section	16.3	32	14.7	33	1.6	2.38	<.01
B	K	Primary School Mathematics Criterion Test, Concepts Section	26.9	32	19.3	33	7.6	5.44	<.001

*t was calculated as $\sqrt{F}$.

**Adjusted mean scores obtained in ANCOVA are shown in parentheses; pretest total score was the covariate.

Table 12

Ronshausen's Summary of Math Computational Skills:
Mean Scores for Treatments

School System	Grade	Subtest	Treatments Tutored	N	 Untutored Control	N	Diff. in Means	t	p
A	1	Metropolitan Achievement Tests	9.8 (10.2)[**]	140	10.8 (10.1)	85	.1	.80[*]	NS
A	1	Criterion Items Subtest, Computational Skills Section	4.3 (4.4)	140	4.5 (4.2)	85	.2	.74	NS
B	1	Metropolitan Achievement Tests	13.4	136	13.1	78	.3	.36	NS
B	1	Primary School Mathematics Criterion Test, Computational Skills Section	10.5	136	9.3	78	1.2	2.75	<.01
B	K	Metropolitan Achievement Tests, Primer, Computational Skills Section	9.7	32	4.9	33	4.8	3.81	<.001
B	K	Primary School Mathematics Criterion Test, Computational Skills Section	7.6	32	3.8	33	3.8	4.49	<.001

[*] t was calculated as $\sqrt{F}$.

[**] Adjusted mean scores obtained in ANCOVA are shown in parentheses; pretest total score was the covariate.

Discussion. The extra 15 minutes of instruction in the form of individualized tutorial assistance that was received daily by tutees should certainly have been expected to produce gains. The gains achieved, then, even the dramatic ones (such as the kindergarten mean computation score for tutees which was twice the mean for control students, or the kindergarten mean concepts score for tutees which equalled the score of untutored first grade children) are more in line with expectations than were the Ellson results. The difference may be due to the greater sensitivity to instruction of math tests as opposed to reading tests. Even within math tests, different parts may be differentially sensitive to instruction. Ronshausen's results suggest that concept attainment is more sensitive to instruction than computational skills. This was also a finding in the Fitz-Gibbon study: tutors worked on three multiplication tables almost daily with their tutees yet on speed tests of computational fluency, given at posttest time, no significant differences could be detected between experimental and control groups, either for tutors or tutees. This contrasted sharply with the effectiveness of the tutoring in teaching students how to add fractions. Another tutoring study reported the same phenomenon: Zack, Horner and Kaufman (1969) report that, following 30 minutes of tutoring by university students 2 days per week for 6 months, fourth graders at an inner city school showed gains significantly higher than those of a control group in arithmetic *concepts* but no significant differences in arithmetic *computation* (both concepts and computation were measured by the Stanford Achievement test).

Ronshausen (1975). Ronshausen followed her 1974 study with one using juveniles as tutors in place of paraprofessionals. It should be remembered, in assessing this report, that the focus was still on benefits to tutees rather than tutors. Tutor behavior was determined by the programmed materials so that no knowledge of mathematics was required.

Ronshausen's 1975 report described programs in several schools employing children from fourth grade through high school as tutors. Usually the cross-age tutoring program was initiated because there were no more funds for paraprofessionals but some programs were initiated in the hope of positive attitude gains for the tutors.

Ronshausen pointed out several characteristics common to all the projects:

- The tutoring materials and projects were used as written, regardless of ages or assumed abilities of tutors. Tutors were expected to follow the programs word-for-word. [Comment: This might limit the assignment of a tutor's role to students with adequate reading achievement.]
- Tutor training, essential to the proper implementation of programmed tutoring, was done by local, professional personnel, usually the same people who later supervised the tutors' work. [Comment: The use of student-tutors thus reduced the cost of the program, but not to zero. Professional personnel were still required to train and supervise.]
- All project directors agreed that student tutors required more supervision than adult tutors. "The supervisor of peer tutors must pair tutors and tutees, schedule both tutors and tutees, arrange tutor transportation, prepare all tutoring materials, accessories and tutoring stations and

handle all discipline problems. The supervisor of peer tutors may have greater problems in establishing and maintaining rapport with classroom teachers and school administrators than in an adult-tutor situation. It is sometimes difficult for the latter groups to understand that peer tutoring can be effective and they may be even more apprehensive when peer tutors are themselves low achievers or 'discipline problems'" (p. 8). [Comment: Ronshausen has clearly stated here the considerable work associated with running a cross-age tutoring project and has touched on a facet which can make or break a program in which secondary school students work in an elementary school; namely, relations between tutors and the elementary school teachers.]

- All project directors agreed that a comprehensive master plan for the project was essential and had to include in-service training for classroom teachers, administrators and subject matter specialists. "These professionals must have some information about programmed tutoring and why it is effective, how tutors and tutees will be selected, where and when tutoring will be done, why peer tutoring is begun and how the professionals will participate in the project. Other components of the comprehensive master plan include the following: selecting tutors and tutees, tutor training, establishing tutoring stations, providing tutoring materials and incidental supplies, scheduling, providing tutor transportation, supervising tutors, testing (if needed) and evaluating the project" (p. 9). [Comment: The need for laying the groundwork for an innovation, fully explaining it to those involved and keeping everybody informed is quite clear. The allowance made in the recommended implementation strategy for Learning-Tutoring Cycle projects (CSE Report on Tutoring #2) for pupil-free time for the sending teachers is designed to facilitate this communication.]

Summary of Programmed Tutoring

Programmed tutoring has been developed primarily for effective instruction in basic skills for primary grade children. Tutors have usually been paid paraprofessionals. The effectiveness of these empirically developed materials was not questioned but several points were emphasized. First, tutors may not benefit cognitively from following highly prescribed materials and may become bored. A couple of experiments indicating the lack of cognitive gains in the situation or a lack of active cognitive processing were cited ("Thinking," said Dewey, "is the method of an educative experience.").

Secondly, contentions that "directed" tutoring is effective were rejected because of a confounding of tutor experience with the programmed or directed tutoring treatment, the possibility of insensitive tests and the possibility of instruction having been matched to the test in the programmed treatment. The question of the differential sensitivity to instruction of various parts of tests was raised by examination of Ronshausen's results. Ronshausen (1974), Fitz-Gibbon (1975) and Zack, Horner and Kaufman (1969) have all found computational skills to be apparently less easily influenced by instruction than math concepts.

Ronshausen reported a study in which high school students rather than paraprofessionals were tutors and pointed out several consequences of this change for project management.

CHAPTER IV. NON-COGNITIVE OUTCOMES OF TUTORING

It will come as no news to anyone that non-cognitive outcomes are perceived as important but are difficult to measure. Practitioners make constant reference to the improved self-concepts, motivation and attendance of tutors but quantitative evidence is scarce.

One problem is that atypical behavior is salient to the observer but only occasionally makes a significant impact on measures of central "tendency," the measures usually reported and dealt with by statistical tests. For example, the student who was always absent but then suddenly attends school to be a tutor is noticed, but the average attendance might be scarcely influenced statistically. In Fitz-Gibbon's study, a student out of school with his leg in a cast hobbled in for tutoring, and a young lady reported for tutoring early in the morning with her cousin before they cut school for the rest of the day; but overall attendance figures showed no strong effects. High schools have reported discovering that students who were present as tutors for weeks were attending no other classes, but data are rarely collected. There is a human tendency to weight positively the salvation of the one lost sheep but unless this weighting is added to statistical procedures, we shall probably find that tutoring has no effect on attendance.

Criscuolo (1973) reported no improvement in sixth grade tutors' attendance during a 30 week tutoring project. Bremmer (1972) reported that 60 percent of tutors improved their attendance record but without further information, this figure cannot be interpreted as good or poor. Perhaps 80% of other students improved their attandance, for example.

What of student satisfaction? We shall deal first with tutee satisfaction. Most practitioners report enthusiasm from students although this enthusiasm can wear thin in prolonged programs. Tutees have generally been reported keen to participate (Mohan, 1972; Snapp, Oakland & Williams, 1972; Paducah, Kentucky program--Klaus, 1975, p. 57). Fitz-Gibbon found 67 percent of the randomly selected fourth grade tutees would have liked to continue being tutored, 12% were unsure and 21% did not wish to continue. Tutees who reported being given considerable amounts of drill on multiplication tables were significantly less likely to want to continue with tutoring than those reporting less drill work. Blank, Koltuv and Wood (1972) reported a marked resistance of kindergarten tutees to tutorial sessions with paraprofessional adults. As Sherertz has noted, "tutoring" is often associated with a need for special assistance and may be resisted if it is perceived as remedial. An important influence on acceptance of tutoring by tutees will be a consideration of what they are missing by receiving tutoring. This consideration apparently applies to many innovative projects; although the projects are well-liked, students do not like to be pulled out of class to participate because they might miss something (Stearns, 1977).

Turning now to tutor satisfaction, it must be noted first that tutors have been volunteers in the majority of projects reported, and might therefore be expected to like tutoring. As with the attendance question, it is often the extreme cases that are noticed. For example, a report from the Salembier project (Site #87 in CSE Report on Tutoring #3, p. 20) of a student who wanted to do nothing but tutor. Mohan (1972) reported a unanimous wish to continue tutoring, and it is significant that his program involved "unmotivated" students. Many practitioners declare that tutoring is most valuable for students who are doing least well in the regular classroom. For example: "Many of the best tutors are among those who have problems in school themselves, and this includes an occasional tutor with a juvenile delinquency record" (Klaus, 1975, p. 57 writing about the Paducah, Kentucky program). Fitz-Gibbon (1975) found a significant negative correlation between tutors' wishes to have further opportunity to tutor and teachers' ratings of the behavior of the students in the regular classroom: the worse their usual classroom behavior, regardless of ability level, the more keen students were to be tutors (see Figure 10).

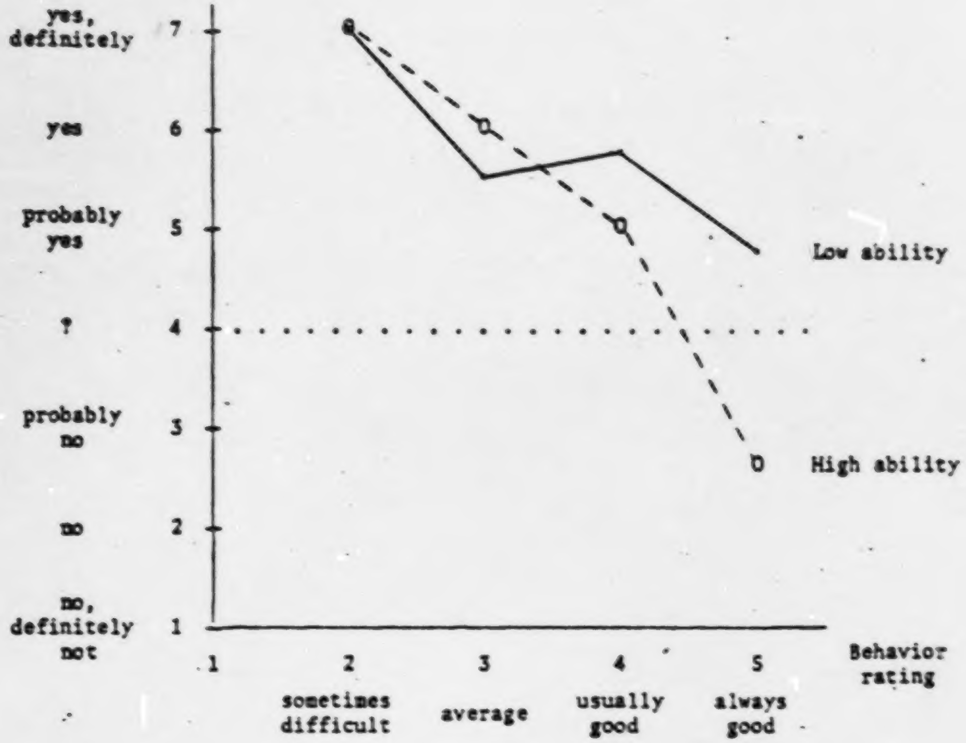

Figure 10. Average responses to the question "Would you like to have tutoring as an elective one quarter?"

The tutors in Fitz-Gibbon's study were not volunteers. Among the 40 randomly selected students who had tutored, 66 percent responded positively to the idea of having tutoring as an elective in the future, 10 percent were uncertain and 24 percent were negative. An important finding three months after the project was that students who had been required to tutor were significantly more likely to wish to tutor in the future than those who had not been required to tutor: non-voluntary exposure to tutoring had increased students' wish to tutor (see Figure 11). (Analysis of variance of the data in Figure A10 showed significant main effects for tutoring and sex.)

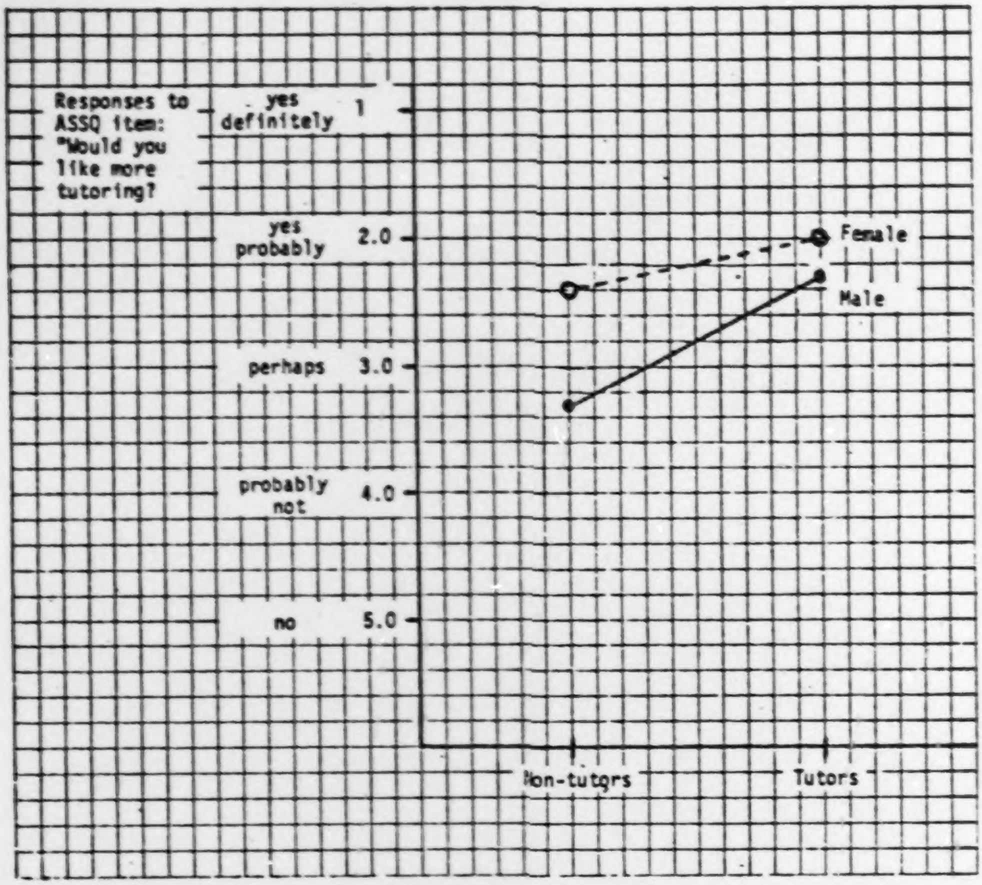

Figure 11. Responses of tutors and non-tutors to tutoring: mean responses of male and female ninth graders who had and had not experienced being a tutor, to question about tutoring in the future.

What determines if a tutor enjoys tutoring and wishes to continue? Students will probably vary widely in their response to the tutoring role but the response will be affected by the relationship established with the tutee. Fitz-Gibbon's data showed significant correlations between tutors' reported enjoyment of tutoring and the extent to which tutees had lived up to tutors' role expectations for them: trying hard to learn the work, and appearing to like the tutor and enjoy the tutoring. Fourth graders are already not totally cooperative in school and tutors did not like trying to persuade or coerce reluctant tutees into working. Ronshausen (1974) reported that high school tutors preferred to work with the primary grades: "...high school tutors prefer to work with first and second graders. The tutors may accept a few fourth graders but they will resign rather than tutor only older children" (Ronshausen, 1975, p. 5).

Gallup polls have consistently reported "discipline" as the major perceived problems in schools, and junior high schools usually represent a high point in adversary relationships between teachers and students. One of the strongest immediate benefits to teachers from LTC programs might be a relief from the role of martinet. These are reports from two tutoring projects implemented in inner-city junior high schools:

> The supervisor of the (seventh grade) tutees...remarked that in September the twenty tutees in each class were very "rambunctious and playful." It was difficult to get them to work in this large group because their need for help was so great. When each student was assigned a tutor in December, 1972, the whole atmosphere of the class changed, and it became a classroom "alive with work." Discipline problems, she noted, became almost non-existent" (Hoffmeister, 1973, p. 26).
>
> For three weeks low achieving ninth graders (tutors) worked hard and presented no discipline hassles. They were colleagues, seeking advice and assistance, not students playing to a peer group and enjoying such diversions as the "she's got my pencil" disruption and the "Don't mess with me" routine (Fitz-Gibbon, 1975(b), p. 3).

When tutors themselves face the problem of trying to motivate students to work, their empathy for teachers might increase, an effect which is predicated from work in social psychology (see CSE Report on Tutoring #6) but which, it appears, no one has attempted to measure. Many practitioners comment on the effect, however, and tutors themselves often mention it spontaneously.

A change in tutors that several projects have tried to assess is in the area of self-concept. Damico and Watson (1974) examining tutoring in a selfcontained, multi-grade classroom are reported to have found positive effects from engaging in peer-helping relationships on self-esteem (Bloom, 1976, p. 41). Robertson (1971) using a semantic differential mean score, found that fifth graders who had tutored first graders three days a week for two months had higher self-concepts than control group students who had either received tutor training but had had tutoring delayed (purportedly by non-availability of tutees) or who had received no treatment. Otherwise, results of attempts to measure self-concept gains have not been positive. Cloward (1969) found no effects on self-concept, nor did Liette (1971) working with Black male underachievers for 12 weeks, nor did Boyd (1969) working with sixth graders tutoring third graders for 8 weeks, nor did Mohan (1972) working with unmotivated sixth graders for 12 weeks.

It is likely that self-concept is too global a measure at which to look for specific effects. As with cognitive outcomes, *the closer the assessment is tied to the actual content of the tutorial sessions, the more likely it is to reflect the effects of tutoring.* Paoni (1971) reported that interest in reading increased among sixth graders who tutored in reading, and Werth (1968) reported more interest in English among tutors than among controls, after tutoring in English. Both these programs ran several months. Fitz-Gibbon (1975), however, was unable to detect any increased liking for mathematics or confidence in mathematics following being tutors in math. However, the questionnaire assessing these attitudes was given three months after a brief three week project. Projects of considerable duration or intensity are probably required for the production of significant attitude changes.

Summary of Non-Cognitive Outcomes

Quantitative findings regarding the effects of tutoring on self-concept, attendance, attitudes to school, and attitudes to school subjects are not strong, but it is widely believed that adequate measurement devices would find them to be positive. It may be that researchers assessing the non-cognitive effects of tutoring interventions will need to focus on subjects in extreme categories to locate significant changes in attitudes and obtain results that are congruent with the rather consistent reports of hundreds of practitioners. Practitioners observe salient cases, but most statistical tests in use examine central tendencies. It is also likely that, as with cognitive effects, the affective impact of tutoring will be closely related to the tutoring process, not a generalized effect on such a complex attitude as "self-concept."

It is vital that, difficult though it may be, research into the non-cognitive effects of tutoring projects continues. The effects observed by practitioners (Report #3) and predicted by theory (Report #6) are effects which are important for schools and for society in general.

CLASSIFICATIONS OF TUTORING PROJECTS

The classification employed in these reports distinguishes between projects according to the manner in which the tutoring task is selected--whether based on needs of the teacher, the tutee, or the tutor. Below two other tripartite classifications that have been proposed are discussed.

Sherertz (1975) proposed a classification quite similar to the one adopted in these reports. Sherertz, a developer of the Ontario-Montclair "Cross-age teaching" projects in which considerable emphasis was placed on the helping relationship, classified tutoring projects as: Aide projects in which older students assist teachers, Tutoring projects in which somebody --perhaps another student or a paid paraprofessional or adult volunteer--works on a one-to-one basis to help the tutee, and Cross-age teaching programs in which older students help younger students to the benefit of both in academic and social areas.

The first two categories correspond to Teacher Aide and Tutorial Service projects, but the third category adopts both cognitive and affective goals for both tutors and tutees. Can one necessarily plan for all four goals simultaneously, without recognizing any priorities? The need for priorities to be established cannot, we suggest, be avoided. For example, if non-cognitive (affective) goals for tutors were of primary concern, the project should be developed to most effectively reach these goals, and this might entail neglecting cognitive goals for tutors. It might be found, for example, that desirable, responsible behavior and a caring relationship was best evoked in tutors when they were given kindergarten children to tutor rather than, say, fourth graders. It is not likely that secondary students would learn much in the way of cognitive skills in a project in which they worked with kindergarten children, but such a project might be the most effective in achieving non-cognitive goals.

The last category adopted by Sherertz seems to imply that all four goals are of equal importance or to presuppose that they occur together. The example just presented indicates that choices must sometimes be made on the basis of goal priorities. Can it be assumed, nevertheless, that positive outcomes are generally correlated? For example, are projects that students like usually effective academically? While it is often assumed that a project that promotes students' enjoyment of school will be cognitively successful, such a link between affective and cognitive outcomes has not been established.

Attempts to measure liking and disliking of school and to relate these attitudes to school success have shown that students may like school but do poorly, as well as like school and do well. Similarly, from studies of instructional television (ITV), we find that "liking ITV is not always correlated with learning from it" (Chu and Schramm, 1967, p. 123). Studies of Follow Through classrooms found no significant correlations between children's negative affect and math scores, with entering ability controlled for statistically (Stallings and Kaskowitz, 1975, p. 281).

These few examples serve as a warning that there may be an independence between cognitive and affective outcomes, thus implying that cognitive and affective goals may require different strategies. (There is a kind of moral corollary to this independence: If cognitive and affective outcomes are independent, it forces a consideration as to the relative value and importance of each in their own right. For example, if project planners are seeking student satisfaction and enjoyment, this kind of outcome may have to be valued for itself, rather than defended as a way to promote cognitive achievement.)

Although studies of tutoring projects may find that some outcomes are correlated (for example, anecdotal evidence and observation indicate that tutors are not pleased with their role if tutees do not appear to be learning), in the meantime it seems cautious to avoid making assumptions regarding the possibility of attaining all goals simultaneously. The classification adopted, therefore, assumed that one might plan for cognitive or affective or other goals for tutors, tutees or teachers, and classifies projects according to the project planning procedure which indicates the "target group."

If projects may be designed primarily to serve tutees, it seems acceptable to suggest that they might also be designed primarily to serve tutors. This is not to deny that tutors often benefit in Tutorial Service projects and tutees in Learning-by-Tutoring projects. The question is one of emphasis, priorities and planning procedures, as was illustrated in Figure 1.

Klaus (1975a) proposed that projects be classified by their goals: those aiming at tutor cognitive growth, those aiming at tutee cognitive growth, and those aiming at positive affective outcomes for both tutors and tutees. The "both" classification raises the same problems applied to Sherertz's "all four" classification. Is it possible to plan for both or must one goal take precedence? Do positive affective outcomes come linked together for tutors and tutees? And would a project aiming at positive affective outcomes be noncognitive? If it were found that in tutors the best affective outcomes derived from successfully implemented projects with cognitive goals, the classification would cease to distinguish.

The main problem with a classification by goals, however, rests with the problem of how one knows a project's goals. Does one find out by asking project personnel? Since innovators are usually, of necessity, hopeful and optimistic, a claim to be working towards all goals is not unusual, but it does not provide any information about the nature of the project. Furthermore, there's "many a slip 'twixt cup and lip." A project might have a stated goal of tutor-growth-in-responsibility, but if no attention was paid in project implementation to promoting such growth (such as by increasing tutors' significant responsibilities), the stated goal would be misleading and could scarcely be used to classify the project in a way that is informative regarding the outcomes that might be expected.

The omission by Klaus of Teacher Aide projects could be justified by arguing either that such projects are not really tutoring projects or that they are designed to help the tutee. However, Teacher Aide projects are frequently called tutoring projects, and yet are quite distinguishable from a Tutorial Service project in which tutors are consistently involved in working with individual tutees. Again, the fact that both the children in the classrooms that have an aide, and the older students acting as the aides, benefit in various ways, does not or should not preclude recognition that in a Teacher Aide project the activities of the "tutors" (aides) are chosen to best assist the teacher. Furthermore, assisting the teacher should not be seen as a goal less justifiable than assisting tutees. Rendering assistance to teachers leaves them free to better help their pupils.

Figure 12 contrasts the classification proposed by Sherertz, Klaus and Fitz-Gibbon.

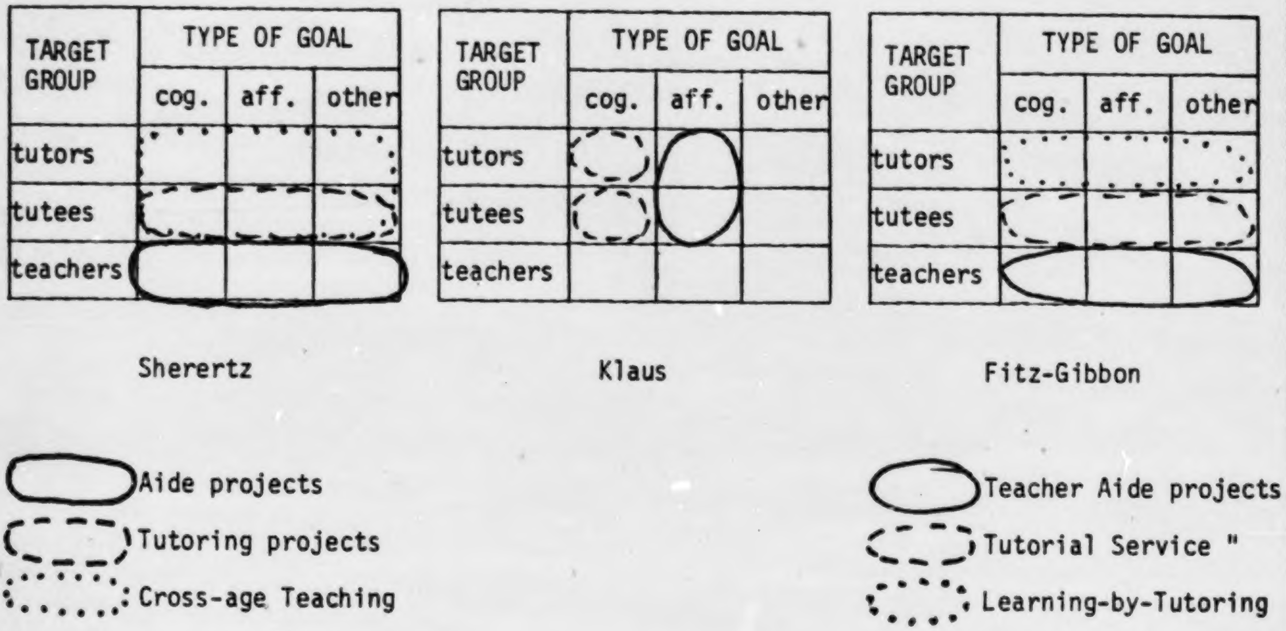
TARGET GROUP
TYPE OF GOAL
cog.
aff.
other
tutors
tutees
teachers
Sherertz
Klaus
Fitz-Gibbon
Aide projects
Tutoring projects
Cross-age Teaching
Teacher Aide projects
Tutorial Service "
Learning-by-Tutoring

Figure 12 Comparison of three classification schemes

The classifications examined are quite gross. As more is learned, by research and evaluation, about methods adopted for working towards various goals, further differentiation of projects will be desirable. Continued focus on the tutoring task--how it is selected, defined and implemented--is recommended.

In this report, programmed tutoring was dealt with separately from less structured tutoring. In general, applying structure to a project means removing options and choices and specifying procedures. The extent to which such structuring is effective for various goals will have to be a topic for formative evaluation in many projects. As a preliminary step in such studies, it should be noted that prescriptions can be made regarding the CONTENT to be taught and/or regarding the METHODS to be employed.

Examples of projects classified by this kind of structuring of the tutoring task are suggested in Figure 13, a "content/method matrix."

METHOD	CONTENT	
	prescribed	open
prescribed	Ellson et al. "Programmed"	Ebersole's reading project
open	Fitz-Gibbon (1975)	Sherertz (Ontario-Montclair project)

Figure 13. Examples of projects classified by the amount of structuring of the tutoring task

In programmed tutoring, both instructional method and content are prescribed. In Ebersole's reading project a method is taught to tutors which they can then employ in working with any reading material. In Fitz-Gibbon's study, objectives for instruction were specified, but tutors could decide daily what activities to undertake with tutees in order to reach the prescribed objectives. In the Ontario-Montclair projects developed by Sherertz and others, tutors could frequently decide what to teach and how to teach it, i.e., both method and content were "open."

Reading About Tutoring

The following CSE Reports on Tutoring are available from the Center for the Study of Evaluation, UCLA Graduate School of Education, 145 Moore Hall, Los Angeles, California 90024.

A. The Learning-Tutoring Cycle: An Overview

An overview of the entire project and recommendations for actions which are designed to put the ideas of this study to the test--to see if the Learning-Tutoring Cycle can indeed significantly improve the educational attainment of disadvantaged students.

#1. Setting up and Evaluating Tutoring Projects

A listing and explanation of decisions that must be made at each school site when a tutoring project is started. Pros and cons for each decision are presented for discussion. A step-by-step outline of evaluation activities is included.

#2. Tutoring and ESEA Title I

A guide to setting up a tutoring project that conforms to Title I requirements and benefits both secondary and elementary school students. Of particular interest to Title I planners, parents, teachers and district personnel.

#3. A Survey of Tutoring Projects

A nationwide survey of tutoring projects and a description of site visits to some existing tutoring projects.

#4. Tutoring: Some New Ideas

Description of a specific approach to tutoring in which the focus is on the learning and motivation of the tutors. Included also are ideas on expansion of tutoring to provide significant educational alternatives, as in a School-Within-A-School project. The reactions of teachers to the ideas are documented.

#5. An Analysis of the Literature on Cross-Age Tutoring*

A literature review that examines both laboratory studies and in-school programs to identify important issues in tutoring.

#6. Tutoring and Social Psychology: A Theoretical Analysis

An examination of tutoring from the perspective of social psychology. Presents operationalized hypotheses that researchers might test.

---

*This report you are presently reading

REFERENCES

Allen, V. L., & Feldman, R. S. Learning through tutoring: Low achieving children as tutors. Journal of Experimental Education, 1973, 42(1), 1-5.

Bausell, R. B., Moody, W. B., & Walzl, P. N. A factorial study of tutoring versus classroom instruction. American Educational Research Journal, 1972, 10(4), 591-598.

Blank, M., Koltuv, M., & Wood, M. Individualized teaching for disadvantaged kindergarten children: A comparison of two methods. The Journal of Special Education, 1972, 6, 207-219.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (Eds.). Taxonomy of educational objectives. Handbook 1: Cognitive domain. New York: David McKay Co., Inc., 1956.

Bloom, S. Peer and cross-age tutoring in the schools: An individualized supplement to group instruction. Chicago: Chicago Board of Education, 1976.

Boyd, G. S. Reading achievement and personal adjustment: A study of the effects of participation as a tutor and as a pupil in an elementary school tutorial reading program. Unpublished doctoral dissertation, University of Alabama, Tuscaloosa, 1969.

Bremmer, B. L. Students helping students program. Seattle: Seattle Public Schools, Planning and Evaluation Department, August, 1972.

Bruner, J. S. Review of A. Gartner, M. Kohler, & F. Riessman, Children teach children. Saturday Review, 1972, 62-62.

Bull, B. L., & Wittrock, M. C. Imagery in the learning of verbal definitions. The British Journal of Educational Psychology, 1973, 43(3), 289-293.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1966.

Chall, J. S. Learning to read: The great debate. New York: McGraw-Hill, 1967.

Chu, G. C., & Schramm, W. Learning from television: What the research says. Stanford, CA: Institute for Communication Research, 1967.

Cloward, R. D. Studies in tutoring. Journal of Experimental Education, 1967, 36(1), 14-25.

Crandall, V. C., Katkovsky, W., & Crandall, V. J. Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations. Child Development, 1965, 36, 91-109.

Criscuolo, N. P. Developing a junior assistants corps in reading. Education, 1973, 93(3), 301.

Damico, S., & Watson, K. Peer helping relationships: An ecological study of an elementary classroom. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April, 1974.

Devin-Sheehan, L., Feldman, R. S., & Allen, V. Research on children tutoring children: A critical review. Review of Educational Research, 1976, 46(3), 355-385.

Dreyer, H. B. Rx for pupil tutoring programs. The Reading Teacher, May, 1973, 810-813.

Ebersole, E. A successful pupil-to-pupil tutoring program. Educational Progress. Los Angeles: Los Angeles City Schools, 1971.

Ellson, D. G., Harris, P., & Barber, L. A field test of programmed and directed tutoring. Reading Research Quarterly, 1968, 3(3), 308-350.

Fitz-Gibbon, C. T. The Role Change Intervention: An experiment in cross-age tutoring. Unpublished doctoral dissertation. University of California, Los Angeles, December, 1975.

Gartner, A., Kohler, M., & Riessman, F. Children teach children. New York: Harper and Row, 1971.

Grannick, L. P. Tutor and tutee achievement. In Youth tutoring youth. New York: National Commission on Resources for Youth, Inc., 1968.

Hamblin, J. A., & Hamblin, R. L. On teaching disadvantaged preschoolers to read: A successful experiment. American Educational Research Journal, 1972, 9(2), 209-216.

Hoffmeister, B. Improving skills in the junior high school through peer tutoring. Unpublished masters' thesis. California State University, San Francisco, August, 1973.

Klaus, D. J. Patterns of peer tutoring. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, 1975. (a)

Klaus, D. J. Patterns of peer tutoring. Washington: American Institutes for Research, 1975. (b)

Liette, E. E. Tutoring: Its effects on reading achievement, standard setting and affect-mediating self-evaluation for black male underachievers in reading. Cleveland, Ohio: Case Western Reserve University, Dept. of Education, June, 1971.

Lippitt, P., Lippitt, R., & Eiseman, J. Cross-age helping program. Ann Arbor: Center for Research on Utilization of Scientific Knowledge, 1971.

Lyman, H. B. *Test scores and what they mean*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

Marks, C., Doctorow, M., & Wittrock, M. C. Reading as a generative process. *Journal of Educational Psychology*, 1975, *67*(4), 484-489.

Mehrens, W. A., & Lehmann, I. J. *Standardized tests in education*. New York: Holt, Rinehart, and Winston, 1969.

Merz, F. Der Einfluss der Verbalisierens auf die Leistung bei Intelligenzaufgaben [The effect of verbalization upon performance on intelligence tests]. *Zeitschrift fur Experimentelle and Angewandte Psychologie*, 1969, *16*(1), 114-137.

Mohan, M. *Peer tutoring as a technique for teaching the unmotivated*. Fredonia, NY: State University of New York, Teacher Research Center, January, 1972.

Myers, K. E., Travers, R. M. W., & Sanford, M. E. Learning and reinforcement in student pairs. *Journal of Educational Psychology*, 1965, *56*(2), 67-72.

Newmark, G., & Melaragno, R. J. *Tutorial community project: Report of the first year*. Santa Monica: Systems Development Corporation, 1969.

Niedermeyer, F. C., & Ellis, P. Remedial reading instruction by trained pupil tutors. *Elementary School Journal*, 1971, *7*, 400-405.

Olds, D. *Cross-age tutoring and parent involvement*. Draft of dissertation study, Cornell University, January, 1976.

Paoni, F. J. *Reciprocal effects of sixth graders tutoring third graders in reading*. Unpublished doctoral dissertation, Oregon State University. Corvallis, 1971.

Raven, J. C. *Standard progressive matrices*. London: H. K. Lewis, 1958.

Robertson, D. J. *The effects of an inter-grade tutoring experience on tutor self-concept*. Paper presented at the Annual Meeting of the California Educational Research Association, San Diego, April, 1971.

Ronshausen, N. L. *The effects of mathematics achievement of programmed tutoring as a method of individualized one-to-one instruction*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 1974.

Ronshausen, N. L. *The programmed math tutorial--paraprofessionals provide one-to-one instruction in primary school mathematics*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, 1975.

Rosenshine, B., & Furst, N. *Summary of the effects of tutoring upon pupil achievement: A review of research*. Paper presented at the Annual Meeting of the American Educational Research Association, 1969.

Sherertz, D. *Ontario-Montclair School District evaluation summary: Cross-age teaching*. Ontario, CA: Ontario-Montclair School District, 1970.

Sherertz, D. Personal communication. La Verne College, California, September 9, 1975.

Snapp, M., Oakland, T., & Williams, F. C. A study of individualizing instruction by using elementary school children as tutors. *Journal of Social Psychology*, 1972, *10*(1), 1-8.

Soar, R., & Soar, R. An empirical analysis of selected follow through programs: An example of a process approach to evaluation. In *Seventy-First Yearbook of the National Society for the Study of Education, Part II*. Chicago: University of Chicago Press, 1972.

Stallings, J. A., & Kaskowitz, D. H. *Follow Through classroom observation evaluation 1972-1973*. Menlo Park, CA: Stanford Research Institute, 1974.

Stearns, M. S. *Evaluation of the field test of project information packages, Vol. 1, Summary Report*. Menlo Park, CA: Stanford Research Institute, April, 1977.

Werth, T. G. *An assessment of the reciprocal effect of high school senior low achievers tutoring freshman low achievers in English classes*. Unpublished doctoral dissertation, Oregon State University, Corvallis, 1968.

Zack, L., Horner, V., & Kaufman, J. Tutoring in a slum school. *Elementary School Journal*, 1969, *70*, 20-27.